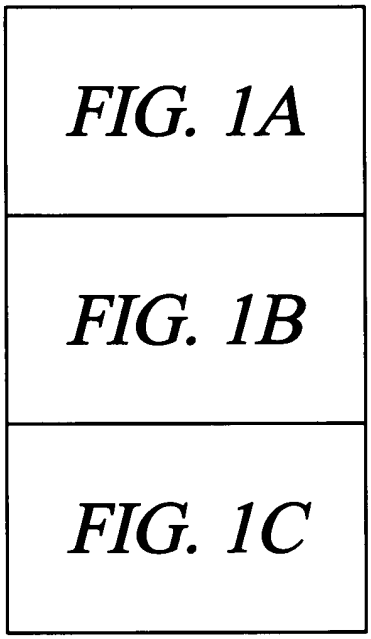


FIG. 1

Anopheles gambiae odorant receptor 1 genomic sequence (SEQ ID NO: 9)

Features:
1) Presumed Untranslated 5' and 3' regions are underlined.
2) Potential TATA box transcription initiation signal is double underlined.
3) Putative Start (ATG) and Stop (TAA) codons are in **BOLD**.
4) Introns are tentatively assigned and are shown in lower case.
Exons are boxed.

AGCTTTGTTCATTTATGTTGAAATCTAGCCCATTTTGTATAGTGCTGAACGACGAAGAACATACGAAAGTACCTCGT
CCGAACACTATCAACATTAATTATACCAAGCTAGAAGAAGATATTTATAGTCAAGCCTCAACATCATAGGAAACTTT
AGCAAAACCATTTAATTTACATGATGATAAGTCCCACCTCTTACCCCAGCACAGGTTTGAGAAGGACGAAAGTATCT
TTACGATAATATTACTCTAAGGTAGTTTTTGAATAAAATAAAAATTTACGTGCAAGTGGTGGCATCGGACATCATTC
GAAAGAATCTACTAAGTCATACACACACCCAAGACGACCGACGTAGTTTCATCTAGAAAAAACGGGTCAGCTCCATC
GAACACGTCAGGACATAACTGCGACATGCGTATGGTCAGTTCCACTAGTGCCAACACTGGTTCCAGGGCACTACCTT
CCGAAGCAGTAGAACCTAATGTATTGGAAATTATTAGGACATACTGCAACATGCATATGGCTAGTTCCGCTGGTACC
AACGATGGCACCAGGACACTATCTGCGGCCTTGTAAAATCACTGTAAAATCTATACAAAAACGGCTTTACCCATACT
TTATCACAAAAACGGCAGGTGAGGGCTGGATTGCTTCAAAGCATTAGAAATATATAATTTCAAAGTCCATAATCTCC
TTAAAAGATAGACAaCAGTAGAGAACACATTTAGTGCTCTTTTCGTTCGAGTTAGTTGCCTTCTCAAGTAAGCGTTT
AATGCTCAATTGTTGTAGATTCGTTGGATGACTCTCGCTACGTGCTATAGTGGTCAATACTTCCAATTAGATTTCAT
AATTAGTTTCCAATTGTCCACGGAAAACCCaCAAAAGAAAAAAAAACTTGTATCTAGGGTGGAATTTTTCGAGAACA
ATTGGACACTTCAT**ATG**AAAAAGGACAGCTTTTTCAAAATGTTAAATAAACACCGTTGGATCCTTTgttggatttca
attctccaaattctgcagaataattctgcaaattttacaaaactgctcaaccaccaataattccaattaatcatctg

FIG. 1A

aacatttaaaactgataattaagatgagtaattgcttcgtcatcacctaagaaatcgattagtttggataaaaagaa
caaattgaaatacaataaagtccctgaattttattcgaataacggcttgaactcatttatttcaaaaacctttgaga
aattcctcgttgaaaattggtctcctatagttctgctaacgggccacttcaaaagcaagaactaacaaaatcataat
tatggtgcaagtaactatcagtaccagtaatcgccattaaaaactttcctcaatttgcggctcgttaccggctaaa
tacagagcagagtaacgggaagtgatcaacgtcgctattagtataacgaggaacgccctccgaaggtgtgttgaagg
acctttcaaattgaaaccaagtactgtttccagttttaaattggatagttataaaatgagccgttcaacgatcggg
catcatttgagtttcatcttcgaggagaaatagatcagtgccactgtttaaccgaaagtaatgaagctgaacaaact
gaacccacggtgggatgcgtacgatcgacgggattcgttctggttgcagttgctttgtttgaaatatttagGCCTAT
GGCCACCGGAAGATACGGATCAGGCAACGCGGAACCGGTACATCGCGTACGGTTGGGCTTTGCGGATCATGTTTCTA
CATCTGTACGCTCTAACGCAAGCCCTATACTTCAAGgATGTGAAGGATATTAATgtgagtctctagttagctattag
tgttccacctgtccataatctgtcttttattgggtagGACATCGCAAATGCATTGTTCGTGCTTATGACTCAAGTGA
CGTTGATCTACAAGCTGGAAAAGTTTAACTACAACATCGCACGGATTCAGGCTTGTCTGCGCAAGCTTAACTGCACA
CTGTATCACCCGAAACAGCGCGAAGAATTCAGgtaagcctgctgggaaatatgactaaaaagagtgctaacaaacga
ctctcctccaaatgtagCCCCGTTTTACAATCGATGAGTGGAGTGTTTTGGCTGATGATCTTTCTCATGTTTGTGGC
TATCTTCACCATCATCATGTGGGTTATGTCGCCAGCCTTCGACAATGAACGTCGTCTGCCcGTGCCGGCCTGGTTCC
CGGTGGACTATCACCATTCGGACATAGTGTACGGTGTACTGTTCCTGTATCAAACCATTGGAATCGTCATGAGCGCA
ACGTACAACTTCTCGACCGATACCATGTTTTCCGGCTTGATGCTACACATAAATGGACAAATTGTGCGGCTTGGTAG
TATGGTTAAAAAGgtgagttacggcgactacttgcctccagtaaggacagggagtttgtttccgttatgatatcatt
ttatcagCTTGGACATGACGTCCCTCCCGAACGCCAATTGGTCGCAACGGATGCGGAATGGAAAGAGATGCGAAAGC
GCATCGACCATCACTCCAAAGTGTACGGTACGATGTACGCTAAAGTAACGGAGTGTGTGCTGTTTCACAAGGACATC
TTAAGgtacgaattgggccaattaattgtgtcatttaaaaagcttgacccaacttttcacagcttcggcgatgaagt
gcaggacattttccaagGATCTATCTTCGCGCAAGTATGCGCGTCTGTAATTATCATTTGTATGACACTGCTGCAAc
TACCGGGGgCGATGTTACGATgGCCGATCTGCTGGGCTGTGGGGTCTATTTGCTAGTAAaGACATCGCAAGTGTTTA

FIG. 1B

TTTTCTGTTACGTAGGGAATGAAATCTCCTATACGgtaggttggacacgtagaggaattaaatgtttgggaagaata
tcaataccaaatagtatgatgtttcgttacagACGGATAAATTTACAGAGTTTGTTGGGTTTTCCAACTACTTCAAG
TTCGATAAGCGTACCAGCCAAGCAATGATATTTTTTCTGCAAATgtgagatagcggtgtatttgtgcagtcagtaca
ttaaatacgttctctatttcagGACTCTTAAAGATGTTCACATCAAGGTGGGAAGTGTCTTGAAGGTTACGCTAAAT
CTTCACACATTTTTGCAGgtatgtaattatgctgtggtatttagcttgaaataagctacaaactttgaaagtaattt
caatctgttttgtagATTATGAAGCTATCGTACTCCTATCTGGCCGTACTTCAGAGCATGGAATCAGAGTAATGGtG
tTAATATCCt**TAA**TGTTGAAATTATATTTTGTTAGATTTATTGCATAAAGTAaTaTTTAATTTTATACATCAAACGT
AAGCCCGCtaGTTTTCAATTAGCCTTTTCCAAAATTTATCAAATTGATTTCGAATTGATTGCAGAGTTTCAGGAATT
TAATCTGATAGGATATCTTGTTTATCCAATAGAGGTGTGGAAGCGTTCCCAAGCCATTCGTTTGATAGTTTATAGCA
CCGTCGAGCAGTTGATCGCTGTGATCGCTAGGCGCACCTGATTTTATCTTTATCTCGCACCTGTTATGGCAAGGGCG
CTTTTCACACGTTTCACACAATATAATGCACATGTATAATGCATTCTTACTTTAGCATTTTTGTTACATATAATACC
AAAATTATGCATTTTTATTCTCACGCAACGATTAGAGGATGACTTcACAAAGGTCCATCTAGTGGTAGGAGGTATAC
AATTATACCTCTCAAAATCTCACAGCAtAATGAGAAACAAAAGGATACCAAGCATACCCTTTTTTTACTTGACAATT
TCATTTGATTTATGTAATAAAGCACTGCaCGTCGACTTCCTAAAA

FIG. 1C

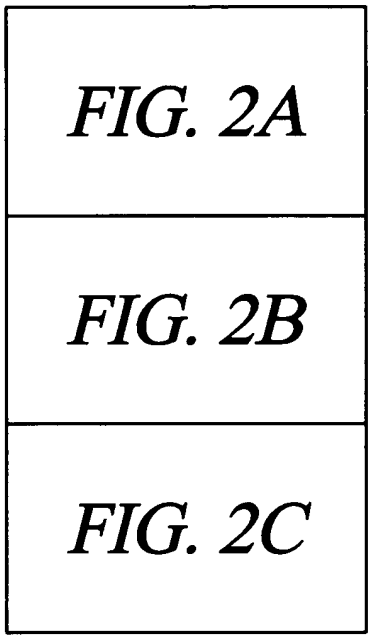


FIG. 2

Anopheles gambiae odorant receptor 2 genomic sequence (SEQ ID NO: 10)

Features:
1) Presumed Untranslated 5' and 3' regions are underlined.
2) Potential TATA box transcription initiation signal is double underlined.
3) Putative Start (ATG) and Stop (TAA) codons are in **BOLD**.
4) Introns are tentatively assigned and are shown in lower case.
5) Exons are boxed.

GGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTCCCTCACCGTGACGTGCTAGAAATGGTTCAACATACTCGT
CCGGCAGAGCGAAGACGACGAACAGCGGAATGTCCCAGGAAATGTAATGAGATATCACAGCAAGTGAACCCAAACCG
AGCTGTGCGCTTTGTGTTGCGCTTTAAAAATGGCCCTTCCTTCGCCGCATCTGCTTGGTTTCACACGCTTTCCCAGG
AAATCCACTGACCACTGGCCACACATCAACCACCGGAGCGGGAGCCTCAGTGCCCAGCGAAGCATATAATTTGCTCA
AAAAGTCACGGTACTCAATTAATTTGATTATAATCAATTTCGTGGCTTCCAACACACCCTTCTTCCACAATCCATCG
CCGAGTGAGCGAGTATAAAGGTGAAGAAACGTACCTTGCGCTTGCTCACTAACTGAACCGGATTTCAAAAAGGAACA
TAAACCGCAACCCACAGCCGAAA**ATG**CTGATCGAAGAGTGTCCGATAATTGGTGTCAATGTGCGAGTGTGGCTGTTC
TGGTCGTATCTGCGGCGGCCGCGGTTGTCCCGCTTTCTGGTCGGCTGCATCCCGGTCGCCGTGCTGAACGTTTTCCA
GTTCCTGAAGCTGTACTCGTCCTGGGGCGACATGAGCGAGCTCATCATCAACGGATACTTTACCGTGCTGTACTTTA
ACCTCGTCGtacgtgggcgaggggaggggcaataaccttcccacttggtggatattttcataccttttccatgtgtt
ttttattctctgtttgttgccatccagCTCCGAACCTCCTTTCTCGTGATCAATCGACGGAAATTTGAGACATTTT
TTGAAGGCGTTGCCGCCGAGTACGCTCTCCTCGAGgtaagtcattggttttctagttttgggggagttgtttaca
ccataaccacccccgacggtaacatttgatcgtcccgcgaaaatgtttgtacagAAAAATGACGACATCCGACCCGT
GCTGGAGCGGTACACACGGCGGGGACGCATGCTATCGATATCGAATCTGTGGCTCGGCGCCTTCATTAGTGCCTGCT
TTGTGACCTATCCTCTGTTTGTGCCCGGGCGCGGCCTACCGTACGGCGTCACGATACCGGGCGTGGACGTGCTGGCC
ACCCCGACCTACCAGGTCGTGTTTGTGCTGCAGGTTTACCTTACCTTCCCCGCCTGCTGCATGTACATCCCGTTCAC

FIG. 2A


CAGCTTCTACGCGACCTGCACGCTGTTTGCGCTCGTCCAGATAGCGGCCCTAAAGCAACGGCTCGGACGCTTGGGGC
GCCACAGCGGCACGATGGCTTCGACCGGACACAGCGCCGGCACACTGTTCGCCGAGCTGAAGGAGTGTCTAAAGTAT
CACAAACAAATCATCCAGtaagtagacgctagtagactcgaccggattgcccttccctcggggaggggaggtttgct
atttcgggatgcggcagcacgcatacacacaaaccggaagccattaattctcccgttttcatgcccgcacgggcact
gggtcatgtttcacatccttccttcctttccaaacacacacacgcgcgcgtgcacgtacagATATGTTCATGATCTC
AACTCACTCGTCACCCATCTGTGTCTGCTGGAGTTCCTGTCGTTCGGGATGATGCTGTGCGCACTGCTGTTTCTGCT
AAGCATTGtaagtaaaatcgaccgacgtgcggtcgctagtccgtctccggactctcatttcgggactcaatcgttcc
atctctcaatagAGCAATCAGCTGGCACAGATGATAATGATTGGATCGTACATCTTCATGATACTCTCGCAGATGTT
TGCCTTCTATTGGCATGCGAACGAGGTACTGGAGCAGGtaatggcgctgaagctgagtttggttgagcggttcgcta
tagatcggctgtcttacattgttgtgtttctgcatggggatcggttttgtttttcctctccatttcagAGCCTAGGC
ATTGGCGATGCCATTTACAATGGAGCGTGGCCGGACTTTGAGGAACCGATAAGGAAACGGTTGATTCTAATTATTGC
ACGTGCTCAGCGACCGATGGTGGTAAGtttggctgatcgatgctctgttcaatgaacatggcacagaaggctgtgta
aatagctgttcattaataagtttttcagaatgtatcgttttagttgatttaaacgcattgttctatgcaatggta
gcaacaatagaccgcctttattaatccaagcttcctttaggattgatttttatttaagagaaagataaaccatttt
tagtaaccaatttagttacaggaaccaaaatacagaatttattattattattattattattattattattatta
ttattattattattattattattattattattattattataattattattattattattattattatt
attattattattaatattattattattattattattactattattattataattattactttattattat
tattattattattattattattattattattattattattattattataattatgattattattatta
ttattattattattattataacaataataattattattatttattattaattaattaatttattattatta
attattattattgttattcattattatacattattatcataataataatttttattatgattattattattatta
ttattattattattattattattattcttattattattattattattattattattaatattatttttaatatt
attattattattattactattcttattataattatttttttttattattattattattattattattattatta
ttattattattattgctattgttattattattcttattattgctattgttattattattattcttattattgttgtt
gttgttgttcttattattgttgttgttgttattcttattattgtttattattattgtttttttttattctctaatta

FIG. 2B

ttccagtaatccataataaaaaataataaagtaaataaatagtaaatagtaaataattccagtaactgtagtaatac
acaataatctctaagaattaaaattgcattttgtaatgaaatatgttgattgttcgaatagttcagaaaaacttaaa
aatgcctcagcattaaacagttttgaggttgttcagggcatttagtttagatattttagtattttaaagcatttgtt
ttcattactacaaaaaagcaaatttatgagtgaattactttcagttcttctaaacgcctatgtgtatgcaattacat
aacaatagctctctttttattgcatttttccttagtaatctaaatccaatctcttctttccctcttgcagATTAAA
GTCGGCAACGTGTACCCGATGACGTTGGAAATGTTTCAAAAATTGCTCAACGTGTCCTACTCCTATTTCACACTGCT
GCGCCGAGTGTACAACTAAACTTAACCGGTAAACAAACAAAAATCCCCTCATCACTATGCAAAGACAGCAAGCAGCC
GATCATCAAACACCATTAGCAGCCACAAAGTTACCAGCCGCTTATCCCACGGGATTTGGTGGAAAGTTATTGCACTG
AAGCTCTTTCACCCAAATTTTCATGGAGGTTCCCTCTCAACCAACCCATTGAAGCGAATAAAAGTATCAGCAACCAG
GCGACGGTGAAAAAACGCTGCATTATTGTGCTTGCTTCAGCATTCCAGCGAATGACTCTTAAACTTTTCCATTCAAA
AGTCGCGATGCTCACGATACGGAGCGGTGTGTTGTTCGATCCGCCGAGTGCACTCGCAAGCCGGTGATGTTGCCGGT
GGAAATGCACAGATCGACACAGCGATAGATAATCGTTTGTTCGCGTAAATGGGAGGGAAAAAAGTAAGCTGCCAGCT
ACTTCATTTCCATGTTAATTGAAACTCAAGCCAACGAACATGCAGAACCCGGTTGGTTGTGTGTCTCCGCTCCGGGA
AAGGTCTCTGCTCCGGGGCATGGATTCTTTCCCCCTCCGGGTGGTTGGGGGTATTGTTTAGGTTTTTATTTTACAAA
TTCATATCCTTCCGCTTCCGCATCAGCCGACCCGGTGGGTGCGCCAGACAGATGTGCGGCGGGCAACAAAACTATGC
ACGAACATGGCCAACAAACACAGCTTCTATCTCATCTCTGTGTCGCACTGTCTCGCTTTCCCGCTGCGTTGCTTGTA
GTACTATCATTGTTTTAGTCCACGGGTTTACTTCTAATTCCATTGCACCACGCAAAAAGGCTCATCCTTTGCTCGTT
CCGGTTGCAACTTCGACAAGCGCATGGTTGGGATACGAACAAAAAACCAACTACTCCACCCACTACTACTACTACTG
CCACCACCACTAACAACACTACACTTGGTTGGGAGCTTGCAGACCCACAAGCAAACAACGATACAAGCTAGCTAGCT
GCTGTGTGCGCTCGAGTCAGCCGACGGTACAAGGTTTAACCGGTACAAGCAACTCCCGGACCGATCCCAAAACTCTG
ACAAGGCACGGGGCCGCATCCGGCAGTACGGTCGGAAAACATGGAAATGTTTAATTAAAACTGTAATTGTCAATCGC
TGCTACAAGTTGTGACACAGGGAGAGAGAGAGACAGAGCGCGCCCGATGGTGATGGTGTAAAAGATAGATACAGGAA
AAGAGCGAGAAACATTGGTACGATTTGGTGTGGTTAGCAAATTTGATTTCCACTGATTTTGAGTGCAAATTTAATGC
ATCGAAAATTTGCCATTCAGGGTAAAGTTGCTCGTGGACGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTAT
CGATACCGTCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGGA

FIG. 2C

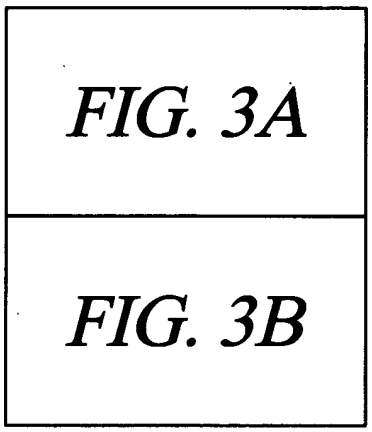


FIG. 3

Anopheles gambiae odorant receptor 3 genomic sequence (SEQ ID NO: 11)

Features:
1) Presumed Untranslated 5' and 3' regions are underlined.
2) Putative Start (ATG) and Stop (TAA) codons are in **BOLD**.
3) Introns are tentatively assigned and are shown in lower case.
4) Exons are boxed.

AAGCAGAACACATCAAGAAGCAATTAGGTGTGTCGTACGTTAGCAAGTAGTTCGCGAGGAGGAATAAAATAG**ATG**CC
TTCTGAGCGGCTTCGTCTCATTACTTCCTTCGGAACTCCTCAAGACAAACGCACGATGGTACTGCCAAAATTAAAGG
ATGAAACAGCAGTGATGCCGTTTCTGCTGCAAATTCAAACCATTGCCGGACTGTGGGGTGACCGTTCCCAGCGGTAC
CGTTTTTATCTCATCTTTTCCTACTTCTGCGCGATGGTGGTTCTACCCAAAGTGCTGTTCGGTTATCCAGATCTCGA
GGTTGCGGTACGCGGCACGGCCGAGCTGATGTTCGAATCGAACGCATTCTTCGGCATGCTAATGTTTTCCTTTCAAC
GCGACAACTACGAGCGATTGGTGCATCAGCTGCAGGATCTGGCAGCTCTAGgtgagtatgcagccaatcgattgttc
caaaccttcgcaacatccttcgtaacactgctacactttcagTCCTCCAAGACCTACCCACAGAGCTGGGAGAGTAC
CTGATCTCAGTGAACCGACGGGTCGATCGGTTCTCCAAAATTTACTGCTGCTGTCACTTTTCCATGGCAACGTTCTT
TTGGTTCATGCCCGTCTGGACGACCTATTCCGCCTACTTTGCTGTGCGCAACAGCACGGAACCGGTCGAGCACGTGT
TGCACCTCGAGGAAGAGCTGTACTTCCTGAACATTCGGACTTCGATGGCGCACTATACGTTTTATGTGGCCATTATG
TGGCCCACGATCTATACGCTCGGGTTTACCGGTGGCACAAAGCTGCTGACCATTTTCAGCAATGTTAAGTACTGTTC
GGCCATGCTGAAGCTCGTTGCACTCCGAATCCACTGTCTAGCGAGAGTAGCGCAAGACCGAGCGGAAAAGGAGCTGA
ACGAGATTATTTCCATGCATCAGCGGGTACTCAAgtaagtaaattcaaattgaaagttttgcagggaataacttgag
tgtgtctgacccgtgcacatcctagCTGCGTGTTCCTGCTGGAGACGACATTCCGCTGGGTATTTTTCGTGCAGTTC
ATTCAGTGTACAATGATCTGGTGCAGTCTCATCCTCTACATAGCGGTGACGgtaatagcattttcgtcatttcgtta
GccttattcaatccatttttgtgaacgtgaatttcccccagGGGTTCAGCTCGACGGTAGCGAATGTATGTGTCCAG
ATCATTTTGGTGACGGTGGAAACTTACGGCTACGGCTACTTCGGAACAGATCTAACCACGGAGGTGCTTTGGgtacc

FIG. 3A

ctttggatgaagcttcaaaaagtaattccaaattctgttttcgatttttccccttttccactagAGCTATGGCGTTG
CCCTCGCCATTTACGATAGCGAGTGGTACAAGTTTTCCATTTCGATGCGCCGCAAACTTCGACTGCTACTGCAACGA
TCCCAAAAACCGCTCGGCGTAACGGCGGGAAAGTTTCGCTTCGTCAATGTGGCCCAGTTTGGCAAGgtaacattaat
tacagtttgaaaattctgaagaatgcatcttacttgccttacttgttgttccagATGCTCAAGATGTCCTATTCATT
TTACGTAGTACTGAAGGAGCAGTTT**TAG**GAGCTGCTGTTTCCCACCCTGGAAATGGCCTTTTCGCACTGTCTTCTGT
TTGTTGGACGCACGCAGCACCGAGAGCGCCCCTGCACGCACTGACGTATTTTGGCTACTTTGACGTTTGCACCTTTG
ACAGCTGAAGGACAGGGTACAATTTTTGCTGCTGTTATTACGCGCAGCGCATTGGATACGAAAACATTGGCCACAAG
TTCTACGATTTTAGCGTTTATTTACTGTTCGTAGCAGCTTTTTTTCCaCAATAAACACACACAATAACGTACCGACAG
TATTCTTTTCATTGTAGGATAGAGAAGCCGCCGGCCAGCAGCCAAAACGCGCCGCAAAACGAAAGGCGGCACCACCG
GGGGAAAAACACGGGAGCAAAACGAGAACAGAACGCAGTAAACAACAAAACCGGCCGGAACAACAACGGTGCCGGAA
ACGA

FIG. 3B

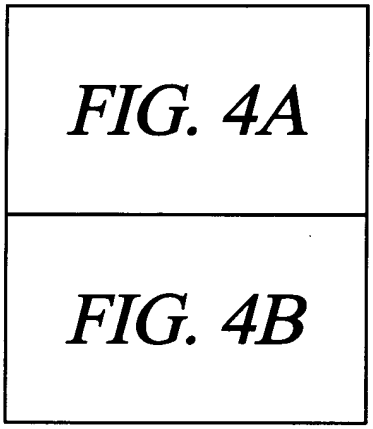


FIG. 4

Anopheles gambiae odorant receptor 4 genomic sequence (SEQ ID NO: 12)

Features:
1) Putative Start (ATG) and Stop (TAA) codons are in **BOLD**.
2) Introns are tentatively assigned and are shown in lower case.

GGGGAACTCCCCCACCCGACCAGACGACGGAAAGCTAACGATGTGCAATTGAATAGTCATTAGT
AGCGTTTTTGCTCGCAAACGAACTAACCCTTTGACTTTTTAAGTTCACTACGGTGAGGACAAAAA
TCAATAAATTAAATCGAGACCGTTGATGAGCAAAAGAAAAAAAAAATATTTTACTGATTTTCATTT
CGTTCCATCGACTACATAATCATAATTATATGCCACATTTTATTATAAGTTTTTGTATCATTTTTA
AACAACACAAAAATGCATCCTTTCGAATATTAGTCAGGTTGTATCAACA**ATG**AAGTTTGAACTGT
TTCAAAAATATTCCTCCCCCGGACACGGTCTTATCCTTCGTGCTAAGGCTTTTGCATATCGTGGGC
ATGAATGGGGCAGGATTTCGGTCGCGAATTCGAGTTGGTGGCATTTTTCTGTTCTATTTAATCTT
TCTTGTAATACCGCCACTAACGGGCGGGTACACCGATGGTCACCAGCGTGTACGCACCAGTGTG
GAATTCCTGTTTAATTGCAATATTTACGGCGGCAGTATGTTCTTTGCCTACGATGTGGCCACTTT
CCAAGCGTTCATCCAGGAACTGAAGAGCCTTTCGGTTTTGGgtaatatttaattaattaaaattgcgtttattgcat
catcatttgtttctctttgcagTATGCTCACATTCGTACAGACTAAAGTATAAGCTGACCCGGTTCAACCGTC
GAGCGGATATTATCGCCAAAGTGCAAACGACCTGCATGGGTGCTGTAACGCTTTTCTACTGGAT
TGCACCGATACCTTCCATCTGTGCGCACTACTACAGGTCGACCAATTCCACCGAACCCGTGCGG
TTTGTGCAACATTTAGAGGTGAAGTTCTATTGGCTCGAGAATCGCACCTCAGTCGAGGACTACAT
AACCTTCGTGCTGATCATGCTACCCGTCGTGGTTATGTGTGGTTACGTATGCAATTTGAAGGTGA
TGACCATCTGCTGCAGCATTGGACACTGTACACTGTACACCAGGATGACTATAGAGATGGTAGA
GCAGTTGGAAAGCATGGCATCAGCGGAACGAACTGCCAGCGCCATACGCAACGTGGGGCAGAT
GCACAGTGGTTTACTGAAATGCATTAGGCTTTTGAACACGTCAATCCGATCGATGCTGATGCTGC

FIG. 4A

AGTGGTTGACCTGCGTGTTAAACTGGAGCATTTCTCTCATCTATCTAACGAACGTGgttagttttgtctt
gtttggaaatccaaaaacaaaaagatggctataattgaactttctattacagGGCATCTCGCTACAATCGGTTACCGTGGT
GGTAATGTTTTTTTCTTGCCACTGCGGAAACTTTCCTGTATTGTTTACTTGGGACGCGGCTTGCGA
CACAACAGCAGCTGCTGGAGCACGCACTCTATGCTACACGGTGGTACAACTACCCAATAGCCTT
TCGCAGCAGCATTAGGATGATGTTGAGACAGTCGCAAAGGCATGCACACATAACGGTGGGGAAG
TTTTTTCGCGTTAATTTGGAAGAATTTAGCAGGATTGTCAACTTATCCTACTCTGCTTACGTCGT
ACTTAAGGATGTAATAAAGATGGATGTACAG**TGA**ATGTTTTTTTTTTTTGGCTTGGCAACGAATGA
AGTTTTCCGAATCTATATTAGATCTAGAATTTAATCTAGATGTCATAATATGATCTTGGCCATGA
CCGGTTCCTGGTTTTGGAACCAATTCTCAAAACAATTTTGAACTTAGGGCGAGGCATGAAATGTC
CCAAGAACCTATCCAAGTTCTGGAACTACATATTACCGAATCTATCCCATTATTGCCTCGGAACT
GGTTTGGTGCTAAATATTTGTCCAAATGTTGGTCCTGGACCTATCCAGACAAAGATCTTCAATTA
TTCCTACCACTGGAACTGATTAATTGATGTAGGAAGTCATGGAGGTGTTCAGGGAGAATTTAAA
CACTAATGTTCCAACTCATTATTTCAAGGGCAATTCTATTTTTTATATGCCCCTACGGATTGATAC
GTATGTATTACTCCATTTCCTGGACTTTGTCTTATTCTTGCTGCTGATTGGACGTGAAATGTTGA
GAAAAAGATTCTTATTTATGAGTGATACAGAGCCTTTAAATACTCCTACGTTGTTTGCTATTTAA
GTATGGCCAGGCTAATCACAATCGCTACTAATGAACAGAATCTCTTCTAATTAAACCCTTTCGAT
TGATAGTGTCAATGTCAATGTCGAGATAATTGAACTGCAAACgATACCTACCTTAAACGGAGCAG
AACACATCAAGAAGCAATTAGGTGTGTCGTACGTTAGCAAGTAGTTCGCGAGGAGGAATAAAAT
AG

FIG. 4B

ANOPHELES GAMBIAE

Preferred DNA Codons

Amino Acids		Preferred Codons
Alanine	Ala A	GCC GCG GCT GCA
Cysteine	Cys C	TGC TGT
Aspartic acid	Asp D	GAC GAT
Glutamic acid	Glu E	GAG GAA
Phenylalanine	Phe F	TTC TTT
Glycine	Gly G	GGC GGT GGA GGG
Histidine	His H	CAC CAT
Isoleucine	Ile I	ATC ATT ATA
Lysine	Lys K	AAG AAA
Leucine	Leu L	CTG CTC TTG CTT CTA TTA
Methionine	Met M	ATG
Asparagine	Asn N	AAC AAT
Proline	Pro P	CCG CCC CCA CCT
Glutamine	Gln Q	CAG CAA
Arginine	Arg R	CGC CGG CGT CGA AGA AGG
Serine	Ser S	TCG AGC TCC AGT TCT TCA
Threonine	Thr T	ACG ACC ACT ACA
Valine	Val V	GTG GTC GTT GTA
Tryptophan	Trp W	TGG
Tyrosine	Tyr Y	TAC TAT

http://www.kazusa.or.jp/codon/cgi-bin/showcodon*(con'd on next line)*
.cgi?species=Anopheles+gambiae+[gbinv]

FIG. 5

Name	SEQ ID NO
Arrestin 1(cDNA)	SEQ ID NO: 1
Arrestin 1(polypeptide)	SEQ ID NO: 2
Odorant Receptor 1(cDNA)	SEQ ID NO: 3
Odorant Receptor 1(polypeptide)	SEQ ID NO: 4
Odorant Receptor 2(cDNA)	SEQ ID NO: 5
Odorant Receptor 2(polypeptide)	SEQ ID NO: 6
Odorant Receptor 3(cDNA)	SEQ ID NO: 7
Odorant Receptor 3(polypeptide)	SEQ ID NO: 8
Odorant Receptor 4(cDNA)	SEQ ID NO: 13
Odorant Receptor 4(polypeptide)	SEQ ID NO: 14
Odorant Receptor 5(cDNA)	SEQ ID NO: 15
Odorant Receptor 5(polypeptide)	SEQ ID NO: 16
Odorant Receptor 6(cDNA)	SEQ ID NO: 17
Odorant Receptor 6(polypeptide)	SEQ ID NO: 18
Odorant Receptor 7(cDNA)	SEQ ID NO: 19
Odorant Receptor 7(polypeptide)	SEQ ID NO: 20

FIG. 6

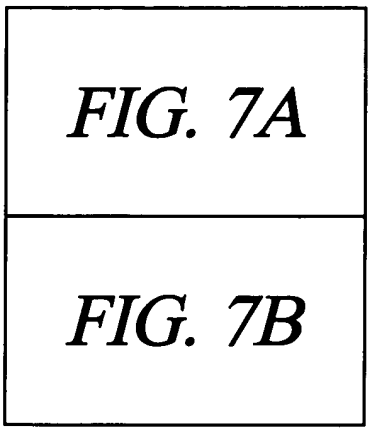


FIG. 7

Anopheles gambiae odorant receptor 5 genomic sequence (SEQ ID NO: 21)

Predicted Exons: *ITALICIZED*, UNDERLINED AND BOXED.
Introns: lowercase.

tctagacttgaacccatgacgggcattttattgagtcgttcgagttgacgactgtaccacgggaccacccgtttatcactatcactatt
aattaattataatatgcttttgtagcgatcagcctaccgggttttgtttctctggatatcttaagttcccatttgattatcaagatagaa
caacaacttgtaccttaaataatcattacgtacccttaatcaacctgtgcatcaaggagttttcgcgaaagcaaaaatccgattgtct
gatgttgtcttgattccatccgattcgttactggttctgcaaaatcgtccaataatacggcaatgtccttatcgatgcttgaatcaacat
cacattgtttgcatttcgtttttgcgtgcaaatatgttatttgcaaagaaggcaaggtaatgtgcttaagagtaaatacaattcgctg
tccatttttgtccaccagtgtgccagaacccgtgcttttagtccttcgaatacatccgaccagtcagcaagcaagtgcatc*ATGG*
TGCTACCGAAGCTGTCCGAACCGTACGCCGTGATGCCGCTTCTACTACGCCTGCAGCG
TTTCGTTGGGCTGTGGGGTGAACGACGCTATCGCTACAAGTTCCGGTTGGCATTTTTA
AGCTTCTGTCTGCTAGTAGTTATTCCGAAGGTTGCCTTCGGCTATCCAGATTTAGAGAC
AATGGTTCGCGGAACAGCTGAGCTGATTTTCGAATGGAACGTACTGTTTGGGATGTTG
CTGTTTTCTCTCAAGCTAGACGACTATGATGATCTGGTGTACCGGTACAAGGACATATC
*AAAGATTG*gtgcgtgataatgattgataaaaggaacctttgagcaactcctatccctttcaag*CTTTCCGTAAGGAC*
GTTCCCTCGCAGATGGGCGACTATCTGGTACGCATCAATCATCGTATCGATCGGTTTTC
CAAGATCTACTGCTGCAGCCATCTGTGTTTGGCCATCTTCTACTGGGTGGCTCCTTCGT
CCAGCACCTACCTAGCGTACCTGGGGGCACGAAACAGATCCGTCCCGGTCGAACATGT
GCTACACCTGGAGGAGGAGCTGTACTGGTTTCACACCCGCGTCTCGCTGGTAGATTAC
TCCATATTCACCGCCATCATGCTGCCTACAATCTTTATGCTAGCGTACTTCGGTGGACT
AAAGCTGCTAACCATCTTCAGCAACGTGAAGTACTGTTCGGCAATGCTCAGGCTTGTG

FIG. 7A

GCGATGAGAATCCAGTTCATGGACCGGCTGGACGAGCGCGAAGCGGAAAAGGAACTGA
*TCGAAATCATCGTCATGCATCAGAAGGCGCTAAA*gtaaggtctgccggtatgttgtggatagaatacattt
ctagctgctttcag*ATGTGTGGAGCTGTTGGAAATCATCTTTCGGTGGGTTTTTCTGGGACAG*
*TTCATACAGTGCGTAATGATCTGGTGCAGCTTGGTTCTGTACGTCGCCGTTACG*gtaacta
aaagcactgtagtgatctgtctgccacaccattcactgctgtgtcttgttttgtcactcttcccag*GGTCTCAGCACAAAAG*
CGGCAAACGTGGGTGTACTGTTTATACTGCTAACAGTGGAAACCTACGGATTCTGCTA
*CTTTGGCAGTGATCTTACCTCGGAGGCAAGTTGTTATTCGCTGA*gtttcagttacttttccgttcccc
tctaaccgtaccacttgtaccatttgtttgagacagagcttgagcgtag*CACGTGCTGCGTACGGTAGCCTCTGG*
TATCGCCGTTCGGTTTCGATTCAACGGAAGCTTCGAATGGTACTGCAGCGTGCCCAGA
AACCGGTCGGCATCTCGGCTGGGAAGTTTTGCTTCGTCGACATTGAGCAGTTTGGCAA
*T*gtatggggagaccttccactgtggcaagaaagattttctttattaatgcatcttttaatttacag*ATGGCAAAAACATCA*
*TACTCGTTCTACATCGTTCTGAAGGATCAATTTTAA*aggggaactcccccacccgaccagacgacggaa
agctaacgatgtgcaattgaatagtcattagtagcgttttgctcgcaaacgaactaacccttgacttttaagttcactacggtgag
gacaaaaatcaataaattaaatcgagaccgttgatgagcaaaagaaaaaaaaatattttactgattttcatttcgttccatcgacta
cataatcataattatatgccacattttattataagttttg

FIG. 7B

Anopheles gambiae odorant receptor 6 partial genomic sequence (SEQ ID NO: 22)

These are the predicted last three exons of another candidate *Anopheles gambiae* odorant receptor.

Predicted Exons: *ITALICIZED*, UNDERLINED AND BOXED.
Introns: lowercase.

aacacccatcttatcggcaaaattagtatttaccgtttgaaagcggcttcccttcctggctgtttctcactctctctctctctgtctctctta
ttgatgccgtatgcgccgcgtgctataggctag*TTATGCTTACCGGATGTTGCGATCGCGCACGTGCTTT*
TCCGCATACGCCAGTGCACACTTGATGGCGGTGGTGATGACGTCTGCTGCGCACCGTT
TTCTGCTCGTGAGTCAGACCTTTTCATTTCCTGCAATATCCTGTTTCTTTCCCGACCCC
ACAGACGGTTAGACGGATATATGCTGGTAAAGTTTGTCCTCTTCATGCTGTGCTTTCTG
*ATCGAGCTGCTGATGCTGTGTGCGTACGGTGAGGATATTGTGGAATCG*gtaaggcaccaggc
ggtgatgagcgagtcgcgagtaattgaagcttttgcttttaaaacacatcagag*CCTTGGGGTGATTGATGCCGCT*
TACGGTTGCGAATGGTACCGGGAAGGGTCGGTGGCGTTCCATCGATCCGTGCTGCAAA
TTATACACCGCAGCCAGCAGTCCGTCATACTGACCGCATGGAAAATTTGGCCCATCCAA
*ATGAGTACTTTCAGTCAG*gtgagttgccaattgattgccgtttgcgttaatatttcagtaagagtgcgctctttcccttag
*ATCCTGCAAGCTTCCTGGTCCTACTTTACCCTCCTGAAGACCGTCTACGGGAATAA*gtaa
gcgcgagagagagagagagagcagtatcgttcacccttggatgaatcaatagatttctaatcatgaaccattgaaaaatgaatca
acattttcgctagttgcacaatattgtaccattctatacagcttcaccacgaccaagcgtttgttgcatcaggaccaaacacgtttcga
caagccgcgtcacctgctggc

FIG. 8

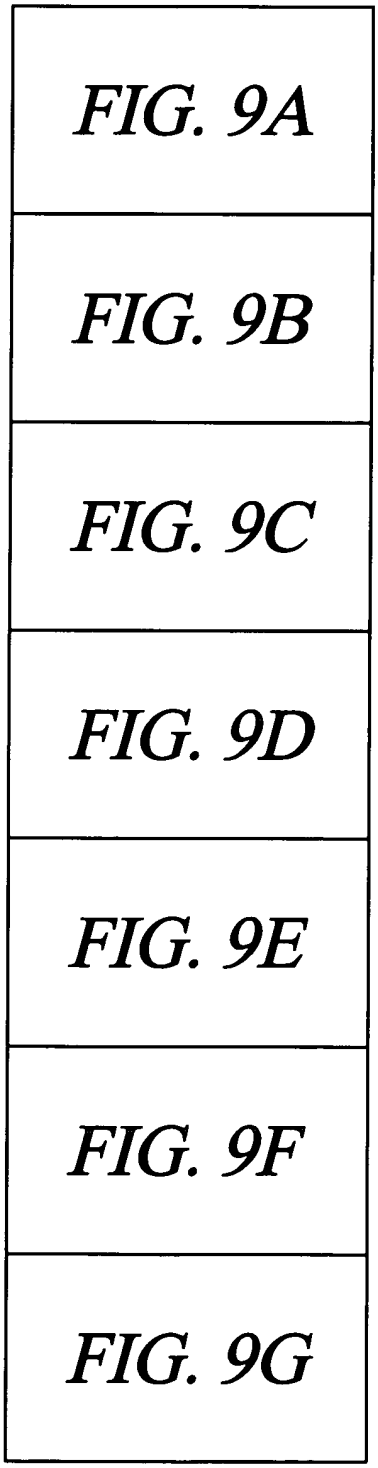


FIG. 9

Anopheles gambiae odorant receptor 7 genomic sequence (SEQ ID NO: 23)

Features
1. Predicted Exons (7): ALL CAPS, ITALICIZED, UNDERLINED, BOXED
2. Introns (6): lowercase
3. 5' and 3' sequences: lowercase, dotted underlined

ccgcccgggcaggtgacttacgcggtctgacttgctggtgcgctgctttgtacggcaaacggctacacaagcgaatcgaattattttcc
tatcacgctgcgcttaccagcgcctgctggtaggcaaagaatgtgcaaagtttcatttggcttggttcgtctgctttgctgtgaacgtgt
gcacggttgcatcgctaaggtttcggtgtgagccgagaagttgcagatcgaaatctctttgtgtgtgtgtgtgtgtgtgcagtgggaa
gcattgtgtttagtgagaagtgaaaagaaaagtgctgaaaaatgcaagtccagccgaccaagtacgtcggccttcgttgccgacct
gatgccgaacattcgggttgatgcaggccagcggtcaactttctgttccggctacgtcaccggcccgatactgatccgcaaggtgtac
tcctggtggacgctcgccc*ATGGTGCTGATCCAGTTCTTCGCCATCCTCGGCAACCTGGCGACGA*
ACGCGGACGACGTGAACGAGCTGACCGCCAACACGATCACGACCCTGTTCTTCACGCA
CTCGGTCACCAAGTTCATCTACTTTGCGGTCAACTCGGAGAACTTCTACCGGACGCTC
GCCATCTGGAACCAGACCAACACGCACCCGCTGTTTGCCGAATCGGACGCCCGGTACC
ATTCGATTGCGCTCGCCAAGATGCGGAAGCTGCTGGTGCTGGTGATGGCCACCACCGT
*CCTGTCGGTTGTCG*gtatgtgtgtatgtgtgtggccgtttgggaaagtgtctttgcggcagaaccccaatctactgttacgc
ttgactgggtttttgttttttctcggtggagggacgggataaaatatctgaaagaataattgagtcaacccacaggggatgcaag
acatcgcaggcagagagtttgggtttgatttatcaccgcacaccgaatatcttcacggttcataagcttcaccgcggtgaaaaggga
actccccatttccctgttttctttttttttcttcctctcgataaattactcatcgcttttcgttttttttttttgttgttgcttctttcttctttcatc

FIG. 9A

cctactag*CCTGGGTTACGATAACATTTTTCGGCGAGAGCGTCAAGACTGTGCTCGATAAG*
GCAACCAACGAGACGTACACGGTGGATATACCCCGGCTGCCCATCAAGTCCTGGTATC
CGTGGAATGCAATGAGCGGACCGGCGTACATTTTCTCTTTCATCTACCAGGTACGTTG
*GCGGAAT*gtcctgcgcgtcacagttggcagtcagtgagcggcaacacggcgaaaaaatgggactaaaaccggtcttcacaga
gccaacacattcctacagcaattgcataccttcgggcggtcgggactgggcaatgcagctacaacatcctcgcctaaagttatgcaat
tcgagcgacaaatgttgccgtgttagggctttttgtgataatagtcgtttttttgtcctctcgcttatcaaactctatcaacggaggaaa
tccattttcgctacaatgcctacagctcaagtttcaaggtcaatcgagcgggtggggatcaacttttttattcattttgctaacgcccca
tcaacaaattctatgttctcaatggcaaagattactgcccgcaccaatcgcccaacgaaacggcaaaagaaaagcgacgattatga
agatgtccaaaccattgcccgcccgacgctttatctgatgatttgcgggatggcttttacttgtctgctactttcaggcacaaaaggaa
atgaaaccagcgcaggctcgtttgccggcttgcggaggttcttcaggcactgaggctgagtacttaaatcgaacgatttttacgattc
tggatccagttttatgatgtggcctgcattacagtggcaattataccctgatgttcatttcattgcattttgtaagtttgtgctggtaacg
cccgtaacgattaattcttttcaaagagattctttcaaagagattcaaaatgtgtataacaaatgctaacgaatggaccgtacttgg
agggttgcggaaagtaacgttttaaaatattcatcacaatcctctgcaaacttgtgcttaattaattggtgcacaataagtttaaact
gtggcggcagatgtgtcgctgtccgcttccttccttcccagcaagctcgtgcgaaataatttattccatcattttaatacagccgtttgtg
cattttaattagcaaagcaatataaaaagcagctaaccatccccattaaaacaaagtgcttccgggcccaattgttatggcggtgga
aagtaatggttttaccagtggaagtgtcctttcccatcgtgggtacttcgcgatattcttgtcttatacaagtgcatacagaaaaaaa
ggacaaatcctccttgctatggtctaaggccagcttcggtaccgcttccgcttcgggatgtcataaagtttgatgggtgtttttaacatt
acttccgctcttaaccacctaatggacttttcatgcttgagctaaagttaaaccagccaccagcggtacgcaccgagccacggttgatt

FIG. 9B

tcggcggcggcctcatccccagttttgcgccaccaatattgccttcattaatctgtaccctcggagcgttagggcccgcggacgagtcct
cgttgtaatgcaccgccatgccacgggacgggataatccgttgggacggcgcgaaagcgactatcgcggacggattggttcgaccg
tgctacaacacattttatgcttcacagatttacttcctgctgttttcgatggtccagagcaacctcgcggatgtcatgttctgctcctggt
tgctgctagcctgcgagcagctgcaacacttgaaggtaggtacggtagcaaacgtggttgtctttacatccgcgtgcagcattatcct
tatcgacgtgtagtgttaacggtaaaagaggaagcgataaaaaagcaacattctctcacaccctcgatctctctttatttttctctctct
ctctctctctctctctctctctctctctctctctctctctctctctctccatctcctcgggcagGGTATTATGCGATCGTTGATG
GAGCTTTCGGCCTCGCTGGACACCTACCGGCCCAACTCTTCGCAACTGTTCCGAGCAA
TTTCAGCCGGTTCCAAATCGGAGCTGATCATCAACGAAGgtatgtgaaacgtgtgctcgtggcagacg
gactcaaagagagcataacacaatcccctggtagttcatttcaatgaccttaacactcggcaagctaagcgagacagtggggacag
tgagaaagagagaacaagaaaaaaaaccatcatccgtacgacatcatcgctacgtaccggtatttcaggatgaggaaataaaac
gctaggggaatgaaagtgcgacagaatgataaaacaatccccacccaggcccccagcctggacgaacggatgtagtgtgcgaagc
gagcaaaaaaagtcaaataaattgaagtttaaaaatagattttccccgtccatccgtggtggagcgtaaagcccggcggacaactt
cgagcacggcgaccgtgcacagtactgtgccacagttgtagggacggataagctccgttccttttttatccttttttttttggagatttgt
ttgcgttcgcatcgttagacgagcttagtgccgtgttgctctaattgctatttattataaagcgcttccaaatagaagatcggttctctc
catttaatctatcgcgcctgtacgcctgaaactatgcactgtgctgtgaaaccgtcaagctcgagcacgacgaatggcccaccgtacc
acgcccgtggtgcccaaagcgcaacgcgaattgcatgttaacaaaccttgcctaccatccaatccgtgtgaaattgcccgctctcttt
ctctcttttgcgctttcggtgtatcgaacggttttgtccctttttttttacttttgctcttgatctcttgctgtgctcactttcatctcatgttttgc
ctgacggtggtgggttttcgaaaaaagagcgatttcttctgcgtgtgtgtgtggttttttaaataaccgctccaggtcgtgttgaacg
ctgcaggaccgatcggagctagtttattatcagctttagtgtttatcccacccatgccccacatcacgtctgtggagagtgggggaag
cttaagtccaatgtaatttaccgtgtttctgtcgttcgtcaccttcttcgtcgatggagattggtgcggttggcacgataaaagcccact
gcacgttacggaccgagggaaaggtctttttgtaggcctagcaacggtcctcattcaccgcatgggggtgtagctcagatggtagag
cgctcgcttagcatgtgagaggtaccgggatcgatacccggcatctccaacccacacaaaacgtttttaagaagattttagggaa
gatattaacgcgggtacactgtgctcctctaagttggaagagtagatgagatgatgacaagggagaaggaacatgtgtacgtgttt

FIG. 9C

gatagcaaacacacaaacaacaatatcatctctgataataatctgatgtgtgatgtgtgtattgttgttatgctgcctttgccatct
tgtccctctctctcctgttcaactcctaaaagaattgtttggagtcctctcagttcctcgtaaagatcctttcgagattcttctttccttttt
attatttattccacgagcctctgacataagtagccttccgcttatttccttctccttgcacttgtcagttccgtgtagagcgtcattttgag
gtttacacatttcccaccgacgcctgattgttacattgtcatctacattgctttccgtttaccgttccgccctttttttttaacgctaccaca
g*AAAAGGATCCGGACGTTAAGGACTTTGATCTGAGCGGCATCTACAGCTCGAAGGCGG*
ACTGGGGCGCCCAGTTCCGTGCGCCGTCGACGCTGCAAACGTTCGACGAGAATGGCAG
GAACGGAAATCCGAACGGGCTTACCCGGAAGCAGGAAATGATGGTGCGCAGCGCCATC
*AAGTACTGGGTCGAGCGGCACAAGCACGTTGTACG*gtaggtatggtaatttctaaggtgtggtgtaaag
cctccaggttccatgaaaaagggatactttaccacagtaagagtttgttttgctggacttacattctttggagcattgtttggtgttgtg
ctgaaaccggttgcaatatcgttttgcgaagaaattatgtgtaaagcgtattacaatctcattcctctgttaatctgtaccaattgtgtc
agccccgaccgaaagcaggcctaattcgtaccagaaaaaccacaagctgtttgtaagcatcgatacgcccgaagctttcaatccagc
caaggcgccacctactattgacgtgacttttgcacgttcacactctccctctcccattctttctataaccaatcgtcgctcagccagcat
cgcccggagtgaagttttatttgaacgatatcacccgtatcgatttccactaaacatgcttaaatcgtttcacaaagctcccccaaa
atcccatttcaccaatccaccaatttgaagtccgtcgtcctttgtgtccttgtgtttgtgtgtttgtgtgagctggagacatgggggagt
gagtaaccgaacaacctcttgccgctgcttcacgatatcgaacagcaccaagataagcatccctttttccctagccgatgtctccgata
tctcgattccgcttccagcgaggcaaagaaaaaggcgaactggctgacctcacccggggcgaggaaaaagcgtagggattacgtc
gagcagcacgagttgtgatttcttcttcttctggttccataaatcgctgacggtttccattaccgcctgcggagtgcacacacgtgaag
ggaaagcgaaaacgtttagattccagcagcaacggcagcaccagaagcagcagcagcgcggcaaattgaatcatcctgacgcgat
gagttgtctgggttttcgggtcggtggcttacagcaccacaccatctgctgcagctaatacagctgtaaatttcgttagacatagactt
gattttacaatattacacacacacttacacacacagctatagatttgtcgcttggcgtatggctctgtacggcgtgccgtacatgccgc
gagccgtgttgctgctggttgcgatacggatcacgtccgattcgattcagcctgcgtgtttttggtgaagatccttatcggtgacccact
ttcagtgtgtcgagagcgagggtcactatggcgcctgtcagttggaaagctaggctcgattcaaagggccattgtgccagtgttcttt
ttaagatagcgataagcttttgatcgaaatagtaaatcaaacattgtttctttttttcctattccaaactgttgccaacctcattattacg

FIG. 9D

ttttgcagcgggtgtatagtaaattgcatactttaaggcgtgattttcaaatgtagcgttccgtatgcagaaacgccatggattatgc
aatttaaacaatgctgcttccttaacattcaaataacggcttattaaggaacttttgtgcaatttgtttttaacagcaaatagttagc
tcagaacgatcacatttagtatcgcttcaacaaagaactcttttaaacacacaatttgtaatgccattccctcgagaaagtttcttgtc
agtcctcctctgcatcacagcaacaaccaaacctgctcatgtttcctgctcgtttcctagctgttttgaacgttatttccgattcctgtgct
tgcccgcttttcttacaatcaaccacaatggttcagatttcgctcttattttattgacccactgctttcgtgctgaagcccgtggaaacaa
tgcgccaagctcagcatccagccatgcatgtaaaatgagccacgcgacagattttagacatcgctttcgctctgcaccggaggtggtt
ttattcttgtttccgattcccacgtccattcgtcctgggtccgtccgccgggcccgaaaccgtaagccgtgcggggaattacgcaatcga
aacgagccagaaaatgagcacgccaaatgcaaagaaaatccccttttgagtggtgctcctgccaccactcatctccccaactggtgg
gtgaaaaaccttgtgcgccccttctctttccagaaaaaaaacgcctcgctcgcacaaaaacatgctcgcccggtgaagctgcgtatgt
cgcagaagctcaaaccaacgccgccagcaagcatcaacaatttctattcaaacacccaacgcagcgcccaaaccgggtgcactgta
ctcagtagcgaagatgctcagattgtcccgtgcgctgctttcgatgcccgtttcggagcgggaagccatcgcttgccaacgttggcgat
gtcttttagccgtggatttgaattttctgaatatcacaggcgggcgcggtttgcctgcaaggttgttgcttcccacacgagcattgcttt
ccgtaccgcggtggggcgagttttcaacgcaaccttctacaagcaacgccacaacgcctgggagcgatatttaacagaaacaagaa
catcccgaacttcagcacatgccgtgatttgcctgttggaaaagcttttgtgagcgtgtgagttgaacgagctctatttcccagcgat
gggtggcatttgtgtggcatgctatcgtcagcttttcttgaatctttacctctccattcgcctccattagtacacgcgtatggaaaatgg
gtgcaacggatcagaacggatttccgcgacagacttaataaagggaaagcaacgcgtttttgcatgtgtagtgtttatgagcttt
atgccgttactttgcaattaaaaatagcaaaaaataacagtttttttttgtaagcggattacaaagaatgtatcagaatattacgtg
aaacattcatttcatgctgttaacgctcaaatagaatagttttgtaacacggattgcataccttgccggtatcggttacattttcgccta
acagtatgcaatctgtttagctttgttgtttaatgactgcgttggtagtacaatatttatttacaccgcgtaatttatctcacaaattgc
aaaaaaatgtcaatctgtatcgattattcacacaaatcagatcccggaaccagtgtagcccaatgtgctcttattgaattaccacga
acaaatcaacctgatgcccgggtccgttggcaaacagcttgcgccgaagccgctcagtgtttcgtgcactaccgtgctgccattttgct
gccctcatcgaacagataaacagaagggcaactcttgtgagcatcgcaatgcccgtctgaagttccgtcgaaaatgggcctaaattc
aatttgacgcatttacccgcgaacaattgcgcgaaggctgtcaagtgtgttccacgaactgcgacaacaagcacacacacaaacac

FIG. 9E

aaatgttatcgtttcggcatgtttctcggtacaaagcgtgtggcgctatgtggcatgccgattcccagacagagtgatcgatagtaaa
tgtagcctatccggtagcattcaatttccttttctatcctcgcaaacaaagcccattctggggaggcgtggtgaagctttcaaaggcat
tgtgaaacaaatgtcctggttcggagggatgctggggaaagcaaacacggtgccgccatcgctgctaccgtcaatcgatcatgcatg
atgtgattaatatttgtgttattcacctgcgtatctatgcgtccgtcgtgtcgttcggatttccggaagtcaaggaaaaagcgactcca
tttgggattggtttttgcagcgaaaaatcaaaacattcgcacaaaaccgtcctccatttcaaatgcctacacttgtcactgtatatctct
ctttctctcgttttgccacgttgcag*TCTCGTTTCAGCAATCGGAGATACGTACGGTCCTGCCCTGCT*
GCTACACATGCTGACCTCCACCATCAAGCTGACGCTGCTCGCCTACCAGGCAACGAAA
ATCGACGGTGTCAACGTGTACGGATTGACCGTAATCGGATATTTGTGCTACGCGTTGG
*CTCAGGTTTTCCTGTTTTGCATCTTTGGCAATCGGCTCATCGAGGAG*gtacgtgcgctcggcgtg
ttgccgtgggaaagcattctccctgccccatatcgcttcattctcccagatcacacatttgcatcacaaagccagcacacttttgcttcg
ccgctgccatctcggcttctgaatgttttcacttctcccatacttctcccgtgcag*AGCTCATCCGTGATGAAGGCGGC*
CTATTCCTGCCACTGGTACGACGGGTCCGAGGAGGCAAAAACCTTCGTCCAGATCGTT
TGTCAGCAGTGCCAGAAGGCGATGACTATTTCCGGAGCCAAGTTTTTCACCGTTTCGC
*TCGATCTGTTTGCTTCG*gtaagtgtagcctggtggctggcacagaacaggctggcaaaacagggactttggctctagc
ctgatgggtggtatatgtgtgtctatttttgctaccattctcgcatcccttcctttccag*GTTCTTGGAGCCGTTGTCAC*
*CTACTTCATGGTGCTGGTGCAGCTGAAGTAA*acagccgtggcccggaaggatgtgtttttttcgctcgttcg
gttgtttgtttgtgcacactttctcttggacattttctctactgcaaaggtttaacaaacagcaacaacaaataatcccaagttttctttt
acagatctttgcaaaatgattagattttaatagattaacagtgcttgattatctgtcctgtagcaaccggggctgaagaacgttgatt
tggtaaaagtacaaaagggacgttggaaattgaaccaccagaagagtgatatttatgcaaagctcaccaagggaaatctatgtat
gtgtgatttgcgctcatcaagcactgtatgtgcctttcaactagtgcagcaataaagagtacaaatgtttcttagcgcaccgtacattg
tcgtttcggcgttttaaccgttgttgataatacacaaaagatgataaaaataaataataacaaaatgttaatatgagtaagtacta
aatagagaaatcgttttagtatgatcatacctccaatcatttgtttgaaattaactttaattttaactcaaattaaaccgatgttttact
ttctgtgagaattattgtggaagaacttaatggaagtataattaaattgattgctaactttatgcgtttttcaatttacgaacgctagt

FIG. 9F

cttcaaacatcgcttcaaaagtattactaccacattattcatttacttatagttatatttattgcctcttcatctttccatggccagaact
actgcagaaaagcttctttttgctcgctttccgatggttggttggacgaagttggtaacaaacggcaagcaattagcataaactatt
ttcgcatcgagatggaaatgaatgtaccactagaaccgagtgaaatgaattacttttcaacttgcacgccaaaaccattatctaaag
tacgcacaacttaaaaacaaaccccaaattgtcgtccacccttcattccactttcttgctacactttccgaccgagttctgtagcgccag
cagcaaaaaaatacatataaaaccttcatcactcaagctgtatcgagccagcgtgggttgtgtttgactgtgctgtgaaagaaaga
agaaaaaaaaaaacacttccacgggaagctagcaattggaaatgcataaattaaccggaagaaattcgcaaaaccccgcaccgac
gtaccgcaccgcatccgtaccgataccggaacaaacggtgtgcgcgaaagaatccgctagcagccccactggcacgggtatttgctt
ttggttctgtgtttttcttccactggtttgggtgcctgggcgaaggctagctcggctactttcccggggccgcaattttctgcagcccaag
gcggcgtgctcgtggggccaaaagaat

FIG. 9G

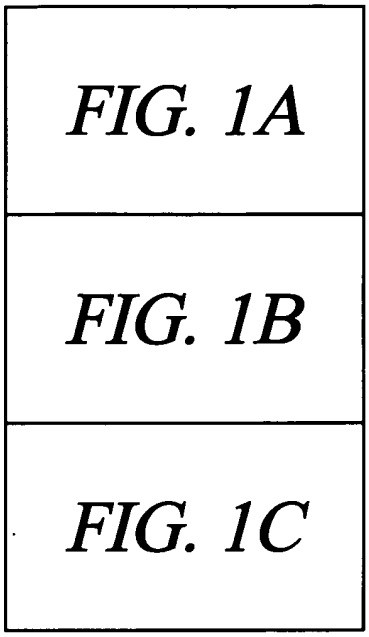


FIG. 1

Now 1A, 1B, 1C due to enlargement.

Anopheles gambiae odorant receptor 1 genomic sequence (SEQ ID NO: 9)

Features:
1) Presumed Untranslated 5' and 3' regions are underlined.
2) Potential TATA box transcription initiation signal is double underlined.
3) Putative Start (ATG) and Stop (TAA) codons are in **BOLD**.
4) Introns ~~are tentatively~~ assigned and are shown in lower case.
Exons are boxed. — was highlighted

AGCTTTGTTCATTTATGTTGAAATCTAGCCCATTTTGTATAGTGCTGAACGACGAAGAACATACGAAAGTACCTCGT
CCGAACACTATCAACATTAATTATACCAAGCTAGAAGAAGATATTTATAGTCAAGCCTCAACATCATAGGAAACTTT
AGCAAAACCATTTAATTTACATGATGATAAGTCCCACCTCTTACCCCAGCACAGGTTTGAGAAGGACGAAAGTATCT
TTACGATAATATTACTCTAAGGTAGTTTTTGAATAAAATAAAAATTTACGTGCAAGTGGTGGCATCGGACATCATTC
GAAAGAATCTACTAAGTCATACACACACCCAAGACGACCGACGTAGTTTCATCTAGAAAAAACGGGTCAGCTCCATC
GAACACGTCAGGACATAACTGCGACATGCGTATGGTCAGTTCCACTAGTGCCAACACTGGTTCCAGGGCACTACCTT
CCGAAGCAGTAGAACCTAATGTATTGGAAATTATTAGGACATACTGCAACATGCATATGGCTAGTTCCGCTGGTACC
AACGATGGCACCAGGACACTATCTGCGGCCTTGTAAAATCACTGTAAAATCTATACAAAAACGGCTTTACCCATACT
TTATCACAAAAACGGCAGGTGAGGGCTGGATTGCTTCAAAGCATTAGAAATATATAATTTCAAAGTCCATAATCTCC
TTAAAAGATAGACAaCAGTAGAGAACACATTTAGTGCTCTTTTCGTTCGAGTTAGTTGCCTTCTCAAGTAAGCGTTT
AATGCTCAATTGTTGTAGATTCGTTGGATGACTCTCGCTACGTGCTATAGTGGTCAATACTTCCAATTAGATTTCAT
AATTAGTTTCCAATTGTCCACGGAAAACCCaCAAAAGAAAAAAAAACTTGTATCTAGGGTGGAATTTTTCGAGAACA
ATTGGACACTTCAT**ATG**AAAAAGGACAGCTTTTTCAAAATGTTAAATAAACACCGTTGGATCCTTTgttggatttca
attctccaaattctgcagaataattctgcaaattttacaaaactgctcaaccaccaataattccaattaatcatctg

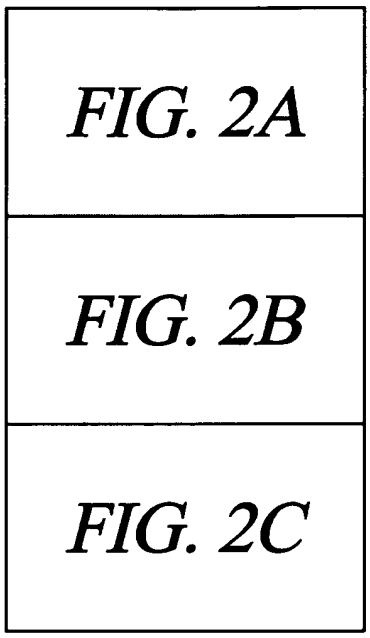
FIG. 1A

aacatttaaaactgataattaagatgagtaattgcttcgtcatcacctaagaaatcgattagtttggataaaaagaa
caaattgaaatacaataaagtccctgaattttattcgaataacggcttgaactcatttatttcaaaaacctttgaga
aattcctcgttgaaaattggtctcctatagttctgctaacgggccacttcaaaagcaagaactaacaaaatcataat
tatggtgcaagtaactatcagtaccagtaatcgccattaaaaactttcctcaatttgcggctcgttaccggctaaa
tacagagcagagtaacgggaagtgatcaacgtcgctattagtataacgaggaacgccctccgaaggtgtgttgaagg
acctttcaaattgaaaccaagtactgtttccagttttaaattggatagttataaaatgagccgttcaacgatcggg
catcatttgagtttcatcttcgaggagaaatagatcagtgccactgtttaaccgaaagtaatgaagctgaacaaact
gaacccacggtgggatgcgtacgatcgacgggattcgttctggttgcagttgctttgtttgaaatatttagGCCTAT
GGCCACCGGAAGATACGGATCAGGCAACGCGGAACCGGTACATCGCGTACGGTTGGGCTTTGCGGATCATGTTTCTA
CATCTGTACGCTCTAACGCAAGCCCTATACTTCAAGgATGTGAAGGATATTAATgtgagtctctagttagctattag
tgttccacctgtccataatctgtcttttattgggtagGACATCGCAAATGCATTGTTCGTGCTTATGACTCAAGTGA
CGTTGATCTACAAGCTGGAAAAGTTTAACTACAACATCGCACGGATTCAGGCTTGTCTGCGCAAGCTTAACTGCACA
CTGTATCACCCGAAACAGCGCGAAGAATTCAGgtaagcctgctgggaaatatgactaaaaagagtgctaacaaacga
ctctcctccaaatgtagCCCCGTTTTACAATCGATGAGTGGAGTGTTTTGGCTGATGATCTTTCTCATGTTTGTGGC
TATCTTCACCATCATCATGTGGGTTATGTCGCCAGCCTTCGACAATGAACGTCGTCTGCCcGTGCCGGCCTGGTTCC
CGGTGGACTATCACCATTCGGACATAGTGTACGGTGTACTGTTCCTGTATCAAACCATTGGAATCGTCATGAGCGCA
ACGTACAACTTCTCGACCGATACCATGTTTTCCGGCTTGATGCTACACATAAATGGACAAATTGTGCGGCTTGGTAG
TATGGTTAAAAAGgtgagttacggcgactacttgcctccagtaaggacagggagtttgtttccgttatgatatcatt
ttatcagCTTGGACATGACGTCCCTCCCGAACGCCAATTGGTCGCAACGGATGCGGAATGGAAAGAGATGCGAAAGC
GCATCGACCATCACTCCAAAGTGTACGGTACGATGTACGCTAAAGTAACGGAGTGTGTGCTGTTTCACAAGGACATC
TTAAGgtacgaattgggccaattaattgtgtcatttaaaaagcttgacccaacttttcacagcttcggcgatgaagt
gcaggacattttccaagGATCTATCTTCGCGCAAGTATGCGCGTCTGTAATTATCATTTGTATGACACTGCTGCAAc
TACCGGGGgCGATGTTACGATgGCCGATCTGCTGGGCTGTGGGGTCTATTTGCTAGTAAaGACATCGCAAGTGTTTA

FIG. 1B

TTTTCTGTTACGTAGGGAATGAAATCTCCTATACGgtaggttggacacgtagaggaattaaatgtttgggaagaata
tcaataccaaatagtatgatgtttcgttacagACGGATAAATTTACAGAGTTTGTTGGGTTTTCCAACTACTTCAAG
TTCGATAAGCGTACCAGCCAAGCAATGATATTTTTTCTGCAAATgtgagatagcggtgtatttgtgcagtcagtaca
ttaaatacgttctctatttcagGACTCTTAAAGATGTTCACATCAAGGTGGGAAGTGTCTTGAAGGTTACGCTAAAT
CTTCACACATTTTTGCAGgtatgtaattatgctgtggtatttagcttgaaataagctacaaactttgaaagtaattt
caatctgttttgtagATTATGAAGCTATCGTACTCCTATCTGGCCGTACTTCAGAGCATGGAATCAGAGTAATGGtG
tTAATATCCt**TAA**TGTTGAAATTATATTTTGTTAGATTTATTGCATAAAGTAaTaTTTAATTTTATACATCAAACGT
AAGCCCGCtaGTTTTCAATTAGCCTTTTCCAAAATTTATCAAATTGATTTCGAATTGATTGCAGAGTTTCAGGAATT
TAATCTGATAGGATATCTTGTTTATCCAATAGAGGTGTGGAAGCGTTCCCAAGCCATTCGTTTGATAGTTTATAGCA
CCGTCGAGCAGTTGATCGCTGTGATCGCTAGGCGCACCTGATTTTATCTTTATCTCGCACCTGTTATGGCAAGGGCG
CTTTTCACACGTTTCACACAATATAATGCACATGTATAATGCATTCTTACTTTAGCATTTTTGTTACATATAATACC
AAAATTATGCATTTTTATTCTCACGCAACGATTAGAGGATGACTTcACAAAGGTCCATCTAGTGGTAGGAGGTATAC
AATTATACCTCTCAAAATCTCACAGCAtAATGAGAAACAAAAGGATACCAAGCATACCCTTTTTTTACTTGACAATT
TCATTTGATTTATGTAATAAAGCACTGCaCGTCGACTTCCTAAAA

FIG. 1C



now 2A, 2B, 2C due to enlargement

FIG. 2

Anopheles gambiae odorant receptor 2 genomic sequence (SEQ ID NO: 10)

Features:
1) Presumed Untranslated 5' and 3' regions are underlined.
2) Potential TATA box transcription initiation signal is double underlined.
3) Putative Start (ATG) and Stop (TAA) codons are in **BOLD**.
4) Introns are tentatively assigned and are shown in lower case.
5) Exons are boxed. — was highlighted

GGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTCCCTCACCGTGACGTGCTAGAAATGGTTCAACATACTCGT
CCGGCAGAGCGAAGACGACGAACAGCGGAATGTCCCAGGAAATGTAATGAGATATCACAGCAAGTGAACCCAAACCG
AGCTGTGCGCTTTGTGTTGCGCTTTAAAAATGGCCCTTCCTTCGCCGCATCTGCTTGGTTTCACACGCTTTCCCAGG
AAATCCACTGACCACTGGCCACACATCAACCACCGGAGCGGGAGCCTCAGTGCCCAGCGAAGCATATAATTTGCTCA
AAAAGTCACGGTACTCAATTAATTTGATTATAATCAATTTCGTGGCTTCCAACACACCCTTCTTCCACAATCCATCG
CCGAGTGAGCGAGTATAAAGGTGAAGAAACGTACCTTGCGCTTGCTCACTAACTGAACCGGATTTCAAAAAGGAACA
TAAACCGCAACCCACAGCCGAAA**ATG**CTGATCGAAGAGTGTCCGATAATTGGTGTCAATGTGCGAGTGTGGCTGTTC
TGGTCGTATCTGCGGCGGCCGCGGTTGTCCCGCTTTCTGGTCGGCTGCATCCCGGTCGCCGTGCTGAACGTTTTCCA
GTTCCTGAAGCTGTACTCGTCCTGGGGCGACATGAGCGAGCTCATCATCAACGGATACTTTACCGTGCTGTACTTTA
ACCTCGTCGtacgtgggcgaggggaggggcaataaccttcccacttggtggatattttcataccttttccatgtgtt
ttttattctctgtttgttgccatccagCTCCGAACCTCCTTTCTCGTGATCAATCGACGGAAATTTGAGACATTTT
TTGAAGGCGTTGCCGCCGAGTACGCTCTCCTCGAGgtaagtcattggttttctagttttgggggagttgtttaca
ccataaccaccccgacggtaacatttgatcgtcccgcgaaaatgtttgtacagAAAAATGACGACATCCGACCCGT
GCTGGAGCGGTACACACGGCGGGGACGCATGCTATCGATATCGAATCTGTGGCTCGGCGCCTTCATTAGTGCCTGCT
TTGTGACCTATCCTCTGTTTGTGCCCGGGCGCGGCCTACCGTACGGCGTCACGATACCGGGCGTGGACGTGCTGGCC
ACCCCGACCTACCAGGTCGTGTTTGTGCTGCAGGTTTACCTTACCTTCCCCGCCTGCTGCATGTACATCCCGTTCAC

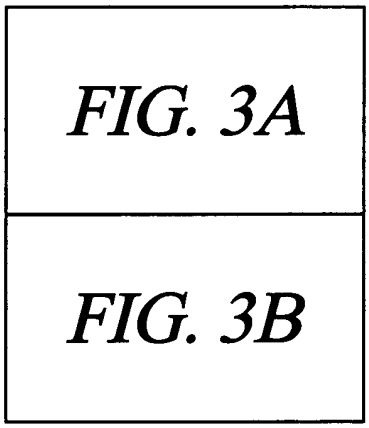
FIG. 2A


CAGCTTCTACGCGACCTGCACGCTGTTTGCGCTCGTCCAGATAGCGGCCCTAAAGCAACGGCTCGGACGCTTGGGGC
GCCACAGCGGCACGATGGCTTCGACCGGACACAGCGCCGGCACACTGTTCGCCGAGCTGAAGGAGTGTCTAAAGTAT
CACAAACAAATCATCCAGtaagtagacgctagtagactcgaccggattgcccttccctcggggaggggaggtttgct
atttcgggatgcggcagcacgcatacacacaaaccggaagccattaattctcccgttttcatgcccgcacgggcact
gggtcatgtttcacatccttccttcctttccaaacacacacacgcgcgcgtgcacgtacagATATGTTCATGATCTC
AACTCACTCGTCACCCATCTGTGTCTGCTGGAGTTCCTGTCGTTCGGGATGATGCTGTGCGCACTGCTGTTTCTGCT
AAGCATTGtaagtaaaatcgaccgacgtgcggtcgctagtccgtctccggactctcatttcgggactcaatcgttcc
atctctcaatagAGCAATCAGCTGGCACAGATGATAATGATTGGATCGTACATCTTCATGATACTCTCGCAGATGTT
TGCCTTCTATTGGCATGCGAACGAGGTACTGGAGCAGGtaatggcgctgaagctgagtttggttgagcggttcgcta
tagatcggctgtcttacattgttgtgtttctgcatggggatcggttttgtttttcctctccatttcagAGCCTAGGC
ATTGGCGATGCCATTTACAATGGAGCGTGGCCGGACTTTGAGGAACCGATAAGGAAACGGTTGATTCTAATTATTGC
ACGTGCTCAGCGACCGATGGTGGTAAGtttggctgatcgatgctctgttcaatgaacatggcacagaaggctgtgta
aatagctgttcattaataagtttttcagaatgtatcgttttagttgatttaaacgcattgttctatgcaatggta
gcaacaatagaccgcctttattaatccaagcttcctttaggattgatttttatttaagagaaagataaaccatttt
tagtaaccaatttagttacaggaaccaaaatacagaatttattattattattattattattattattattatta
ttattattattattattattattattattattattattattataattattattattattattattatt
attattattattaatattattattattattattattactattattataattattactttattattat
tattattattattattattattattattattattattattattattattataattatgattattattatta
ttattattattattattataacaataataattattattatttattattaattaatttattattatta
attattattattgttattcattattatacattattatcataataataattttattatgattattattattatta
ttattattattattattattattattcttattattattattattattattattattaatattattttaatatt
attattattattattactattcttattataattatttttttttattattattattattattattattattatta
ttattattattattgctattgttattattattcttattattgctattgttattattattattcttattattgttgtt
gttgttgttcttattattgttgttgttgttattcttattattgtttattattattgtttttttttattctctaatta

FIG. 2B

ttccagtaatccataataaaaaataataaagtaaataaatagtaaatagtaaataattccagtaactgtagtaatac
acaataatctctaagaattaaaattgcattttgtaatgaaatatgttgattgttcgaatagttcagaaaaacttaaa
aatgcctcagcattaaacagttttgaggttgttcagggcatttagtttagatattttagtattttaaagcatttgtt
ttcattactacaaaaaagcaaatttatgagtgaattactttcagttcttctaaacgcctatgtgtatgcaattacat
aacaatagctctctttttattgcattttccttagtaatctaaatccaatctcttctttccctcttgcagATTAAA
GTCGGCAACGTGTACCCGATGACGTTGGAAATGTTTCAAAAATTGCTCAACGTGTCCTACTCCTATTTCACACTGCT
GCGCCGAGTGTACAACTAAACTTAACCGGTAAACAAACAAAAATCCCCTCATCACTATGCAAAGACAGCAAGCAGCC
GATCATCAAACACCATTAGCAGCCACAAAGTTACCAGCCGCTTATCCCACGGGATTTGGTGGAAAGTTATTGCACTG
AAGCTCTTTCACCCAAATTTTCATGGAGGTTCCCTCTCAACCAACCCATTGAAGCGAATAAAAGTATCAGCAACCAG
GCGACGGTGAAAAAACGCTGCATTATTGTGCTTGCTTCAGCATTCCAGCGAATGACTCTTAAACTTTTCCATTCAAA
AGTCGCGATGCTCACGATACGGAGCGGTGTGTTGTTCGATCCGCCGAGTGCACTCGCAAGCCGGTGATGTTGCCGGT
GGAAATGCACAGATCGACACAGCGATAGATAATCGTTTGTTCGCGTAAATGGGAGGGAAAAAAGTAAGCTGCCAGCT
ACTTCATTTCCATGTTAATTGAAACTCAAGCCAACGAACATGCAGAACCCGGTTGGTTGTGTGTCTCCGCTCCGGGA
AAGGTCTCTGCTCCGGGGCATGGATTCTTTCCCCCTCCGGGTGGTTGGGGGTATTGTTTAGGTTTTTATTTTACAAA
TTCATATCCTTCCGCTTCCGCATCAGCCGACCCGGTGGGTGCGCCAGACAGATGTGCGGCGGGCAACAAAACTATGC
ACGAACATGGCCAACAAACACAGCTTCTATCTCATCTCTGTGTCGCACTGTCTCGCTTTCCCGCTGCGTTGCTTGTA
GTACTATCATTGTTTTAGTCCACGGGTTTACTTCTAATTCCATTGCACCACGCAAAAAGGCTCATCCTTTGCTCGTT
CCGGTTGCAACTTCGACAAGCGCATGGTTGGGATACGAACAAAAAACCAACTACTCCACCCACTACTACTACTACTG
CCACCACCACTAACAACACTACACTTGGTTGGGAGCTTGCAGACCCACAAGCAAACAACGATACAAGCTAGCTAGCT
GCTGTGTGCGCTCGAGTCAGCCGACGGTACAAGGTTTAACCGGTACAAGCAACTCCCGGACCGATCCCAAAACTCTG
ACAAGGCACGGGGCCGCATCCGGCAGTACGGTCGGAAAACATGGAAATGTTTAATTAAAACTGTAATTGTCAATCGC
TGCTACAAGTTGTGACACAGGGAGAGAGAGAGACAGAGCGCGCCCGATGGTGATGGTGTAAAAGATAGATACAGGAA
AAGAGCGAGAAACATTGGTACGATTTGGTGTGGTTAGCAAATTTGATTTCCACTGATTTTGAGTGCAAATTTAATGC
ATCGAAAATTTGCCATTCAGGGTAAAGTTGCTCGTGGACGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTAT
CGATACCGTCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGGA

FIG. 2C



Now 3A, 3B due to enlargement

FIG. 3

Anopheles gambiae odorant receptor 3 genomic sequence (SEQ ID NO: 11)

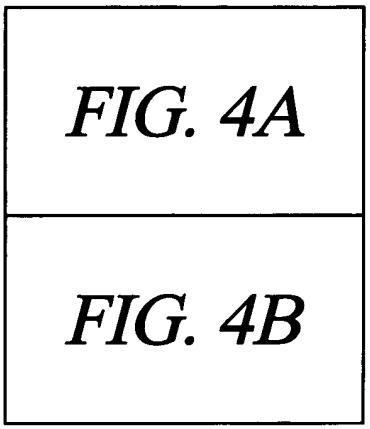
Features:
1) Presumed Untranslated 5' and 3' regions are underlined.
2) Putative Start (ATG) and Stop (TAA) codons are in **BOLD**.
3) Introns are ~~tentatively~~ assigned and are shown in lower case.
4) Exons are boxed. was highlighted

AAGCAGAACACATCAAGAAGCAATTAGGTGTGTCGTACGTTAGCAAGTAGTTCGCGAGGAGGAATAAAATAG**ATG**CC
TTCTGAGCGGCTTCGTCTCATTACTTCCTTCGGAACTCCTCAAGACAAACGCACGATGGTACTGCCAAAATTAAAGG
ATGAAACAGCAGTGATGCCGTTTCTGCTGCAAATTCAAACCATTGCCGGACTGTGGGGTGACCGTTCCCAGCGGTAC
CGTTTTTATCTCATCTTTTCCTACTTCTGCGCGATGGTGGTTCTACCCAAAGTGCTGTTCGGTTATCCAGATCTCGA
GGTTGCGGTACGCGGCACGGCCGAGCTGATGTTCGAATCGAACGCATTCTTCGGCATGCTAATGTTTTCCTTTCAAC
GCGACAACTACGAGCGATTGGTGCATCAGCTGCAGGATCTGGCAGCTCTAGgtgagtatgcagccaatcgattgttc
caaaccttcgcaacatccttcgtaacactgctacactttcagTCCTCCAAGACCTACCCACAGAGCTGGGAGAGTAC
CTGATCTCAGTGAACCGACGGGTCGATCGGTTCTCCAAAATTTACTGCTGCTGTCACTTTTCCATGGCAACGTTCTT
TTGGTTCATGCCCGTCTGGACGACCTATTCCGCCTACTTTGCTGTGCGCAACAGCACGGAACCGGTCGAGCACGTGT
TGCACCTCGAGGAAGAGCTGTACTTCCTGAACATTCGGACTTCGATGGCGCACTATACGTTTTATGTGGCCATTATG
TGGCCCACGATCTATACGCTCGGGTTTACCGGTGGCACAAAGCTGCTGACCATTTTCAGCAATGTTAAGTACTGTTC
GGCCATGCTGAAGCTCGTTGCACTCCGAATCCACTGTCTAGCGAGAGTAGCGCAAGACCGAGCGGAAAAGGAGCTGA
ACGAGATTATTTCCATGCATCAGCGGGTACTCAAgtaagtaaattcaaattgaaagttttgcagggaataacttgag
tgtgtctgacccgtgcacatcctagCTGCGTGTTCCTGCTGGAGACGACATTCCGCTGGGTATTTTTCGTGCAGTTC
ATTCAGTGTACAATGATCTGGTGCAGTCTCATCCTCTACATAGCGGTGACGgtaatagcattttcgtcatttcgtta
GccttattcaatccatttttgtgaacgtgaatttcccccagGGGTTCAGCTCGACGGTAGCGAATGTATGTGTCCAG
ATCATTTTGGTGACGGTGGAAACTTACGGCTACGGCTACTTCGGAACAGATCTAACCACGGAGGTGCTTTGGgtacc

FIG. 3A

ctttggatgaagcttcaaaaagtaattccaaattctgttttcgatttttccccttttccactagAGCTATGGCGTTG
CCCTCGCCATTTACGATAGCGAGTGGTACAAGTTTTCCATTTCGATGCGCCGCAAACTTCGACTGCTACTGCAACGA
TCCCAAAAACCGCTCGGCGTAACGGCGGGAAAGTTTCGCTTCGTCAATGTGGCCCAGTTTGGCAAGgtaacattaat
tacagtttgaaaattctgaagaatgcatcttacttgccttacttgttgttccagATGCTCAAGATGTCCTATTCATT
TTACGTAGTACTGAAGGAGCAGTTT**TAG**GAGCTGCTGTTTCCCACCCTGGAAATGGCCTTTTCGCACTGTCTTCTGT
TTGTTGGACGCACGCAGCACCGAGAGCGCCCCTGCACGCACTGACGTATTTTGGCTACTTTGACGTTTGCACCTTTG
ACAGCTGAAGGACAGGGTACAATTTTTGCTGCTGTTATTACGCGCAGCGCATTGGATACGAAAACATTGGCCACAAG
TTCTACGATTTTAGCGTTTATTTACTGTTCGTAGCAGCTTTTTTCCaCAATAAACACACACAATAACGTACCGACAG
TATTCTTTTCATTGTAGGATAGAGAAGCCGCCGGCCAGCAGCCAAAACGCGCCGCAAAACGAAAGGCGGCACCACCG
GGGGAAAAACACGGGAGCAAAACGAGAACAGAACGCAGTAAACAACAAAACCGGCCGGAACAACAACGGTGCCGGAA
ACGA

FIG. 3B



Now 4A, 4B due to enlargement

FIG. 4

Anopheles gambiae odorant receptor 4 genomic sequence (SEQ ID NO: 12)

Features:
1) Putative Start (ATG) and Stop (TAA) codons are in **BOLD**.
2) Introns are tentatively assigned and are shown in lower case.

GGGGAACTCCCCCACCCGACCAGACGACGGAAAGCTAACGATGTGCAATTGAATAGTCATTAGT
AGCGTTTTTGCTCGCAAACGAACTAACCCTTTGACTTTTTAAGTTCACTACGGTGAGGACAAAAA
TCAATAAATTAAATCGAGACCGTTGATGAGCAAAAGAAAAAAAAATATTTTACTGATTTTCATTT
CGTTCCATCGACTACATAATCATAATTATATGCCACATTTTATTATAAGTTTTTGTATCATTTTTA
AACAACACAAAAATGCATCCTTTCGAATATTAGTCAGGTTGTATCAACA**ATG**AAGTTTGAACTGT
TTCAAAAATATTCCTCCCCGGACACGGTCTTATCCTTCGTGCTAAGGCTTTTGCATATCGTGGGC
ATGAATGGGGCAGGATTTCGGTCGCGAATTCGAGTTGGTGGCATTTTTCTGTTCTATTTAATCTT
TCTTGTAATACCGCCACTAACGGGCGGGTACACCGATGGTCACCAGCGTGTACGCACCAGTGTG
GAATTCCTGTTTAATTGCAATATTTACGGCGGCAGTATGTTCTTTGCCTACGATGTGGCCACTTT
CCAAGCGTTCATCCAGGAACTGAAGAGCCTTTCGGTTTTGGgtaatatttaattaattaaaattgcgtttattgcat
catcatttgtttctctttgcagTATGCTCACATTCGTACAGACTAAAGTATAAGCTGACCCGGTTCAACCGTC
GAGCGGATATTATCGCCAAAGTGCAAACGACCTGCATGGGTGCTGTAACGCTTTTCTACTGGAT
TGCACCGATACCTTCCATCTGTGCGCACTACTACAGGTCGACCAATTCCACCGAACCCGTGCGG
TTTGTGCAACATTTAGAGGTGAAGTTCTATTGGCTCGAGAATCGCACCTCAGTCGAGGACTACAT
AACCTTCGTGCTGATCATGCTACCCGTCGTGGTTATGTGTGGTTACGTATGCAATTTGAAGGTGA
TGACCATCTGCTGCAGCATTGGACACTGTACACTGTACACCAGGATGACTATAGAGATGGTAGA
GCAGTTGGAAAGCATGGCATCAGCGGAACGAACTGCCAGCGCCATACGCAACGTGGGGCAGAT
GCACAGTGGTTTACTGAAATGCATTAGGCTTTTGAACACGTCAATCCGATCGATGCTGATGCTGC

FIG. 4A

AGTGGTTGACCTGCGTGTTAAACTGGAGCATTTCTCTCATCTATCTAACGAACGTGgttagttttgtctt
gtttggaaatccaaaaacaaaaagatggctataattgaactttctattacagGGCATCTCGCTACAATCGGTTACCGTGGT
GGTAATGTTTTTTCTTGCCACTGCGGAAACTTTCCTGTATTGTTTACTTGGGACGCGGCTTGCGA
CACAACAGCAGCTGCTGGAGCACGCACTCTATGCTACACGGTGGTACAACTACCCAATAGCCTT
TCGCAGCAGCATTAGGATGATGTTGAGACAGTCGCAAAGGCATGCACACATAACGGTGGGGAAG
TTTTTTCGCGTTAATTTGGAAGAATTTAGCAGGATTGTCAACTTATCCTACTCTGCTTACGTCGT
ACTTAAGGATGTAATAAAGATGGATGTACAG**TGA**ATGTTTTTTTTTTTTGGCTTGGCAACGAATGA
AGTTTTCCGAATCTATATTAGATCTAGAATTTAATCTAGATGTCATAATATGATCTTGGCCATGA
CCGGTTCCTGGTTTTGGAACCAATTCTCAAAACAATTTTGAACTTAGGGCGAGGCATGAAATGTC
CCAAGAACCTATCCAAGTTCTGGAACTACATATTACCGAATCTATCCCATTATTGCCTCGGAACT
GGTTTGGTGCTAAATATTTGTCCAAATGTTGGTCCTGGACCTATCCAGACAAAGATCTTCAATTA
TTCCTACCACTGGAACTGATTAATTGATGTAGGAAGTCATGGAGGTGTTCAGGGAGAATTTAAA
CACTAATGTTCCAACTCATTATTTCAAGGGCAATTCTATTTTTTATATGCCCCTACGGATTGATAC
GTATGTATTACTCCATTTCCTGGACTTTGTCTTATTCTTGCTGCTGATTGGACGTGAAATGTTGA
GAAAAAGATTCTTATTTATGAGTGATACAGAGCCTTTAAATACTCCTACGTTGTTTGCTATTTAA
GTATGGCCAGGCTAATCACAATCGCTACTAATGAACAGAATCTCTTCTAATTAAACCCTTTCGAT
TGATAGTGTCAATGTCAATGTCGAGATAATTGAACTGCAAACgATACCTACCTTAAACGGAGCAG
AACACATCAAGAAGCAATTAGGTGTGTCGTACGTTAGCAAGTAGTTCGCGAGGAGGAATAAAAT
AG

FIG. 4B

enlarged

ANOPHELES GAMBIAE

Preferred DNA Codons

Amino Acids		Preferred Codons
Alanine	Ala A	GCC GCG GCT GCA
Cysteine	Cys C	TGC TGT
Aspartic acid	Asp D	GAC GAT
Glutamic acid	Glu E	GAG GAA
Phenylalanine	Phe F	TTC TTT
Glycine	Gly G	GGC GGT GGA GGG
Histidine	His H	CAC CAT
Isoleucine	Ile I	ATC ATT ATA
Lysine	Lys K	AAG AAA
Leucine	Leu L	CTG CTC TTG CTT CTA TTA
Methionine	Met M	ATG
Asparagine	Asn N	AAC AAT
Proline	Pro P	CCG CCC CCA CCT
Glutamine	Gln Q	CAG CAA
Arginine	Arg R	CGC CGG CGT CGA AGA AGG
Serine	Ser S	TCG AGC TCC AGT TCT TCA
Threonine	Thr T	ACG ACC ACT ACA
Valine	Val V	GTG GTC GTT GTA
Tryptophan	Trp W	TGG
Tyrosine	Tyr Y	TAC TAT

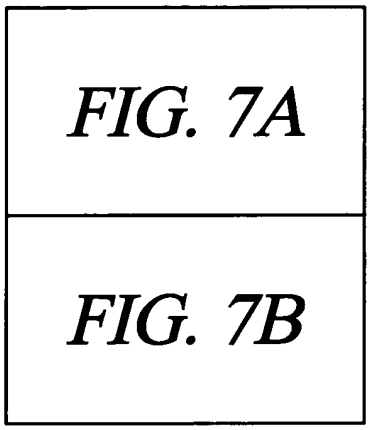
http://www.kazusa.or.jp/codon/cgi-bin/showcodon*(con'd on next line)*
.cgi?species=Anopheles+gambiae+[gbinv]

FIG. 5

enlarged

Name	SEQ ID NO
Arrestin 1(cDNA)	SEQ ID NO: 1
Arrestin 1(polypeptide)	SEQ ID NO: 2
Odorant Receptor 1(cDNA)	SEQ ID NO: 3
Odorant Receptor 1(polypeptide)	SEQ ID NO: 4
Odorant Receptor 2(cDNA)	SEQ ID NO: 5
Odorant Receptor 2(polypeptide)	SEQ ID NO: 6
Odorant Receptor 3(cDNA)	SEQ ID NO: 7
Odorant Receptor 3(polypeptide)	SEQ ID NO: 8
Odorant Receptor 4(cDNA)	SEQ ID NO: 13
Odorant Receptor 4(polypeptide)	SEQ ID NO: 14
Odorant Receptor 5(cDNA)	SEQ ID NO: 15
Odorant Receptor 5(polypeptide)	SEQ ID NO: 16
Odorant Receptor 6(cDNA)	SEQ ID NO: 17
Odorant Receptor 6(polypeptide)	SEQ ID NO: 18
Odorant Receptor 7(cDNA)	SEQ ID NO: 19
Odorant Receptor 7(polypeptide)	SEQ ID NO: 20

FIG. 6



Now 7A, 7B due to enlargement

FIG. 7

Anopheles gambiae odorant receptor 5 genomic sequence (SEQ ID NO: 21)

Predicted Exons: *ITALICIZED*, UNDERLINED AND BOXED
Introns: lowercase.

was highlighted

tctagacttgaacccatgacgggcattttattgagtcgttcgagttgacgactgtaccacgggaccacccgtttatcactatcactatt
aattaattataatatgcttttgtagcgatcagcctaccgggttttgtttctctggatatcttaagttcccatttgattatcaagatagaa
caacaacttgtaccttaaataatcattacgtacccttaatcaacctgtgcatcaaggagttttcgcgaaagcaaaaatccgattgtct
gatgttgtcttgattccatccgattcgttactggttctgcaaaatcgtccaataatacggcaatgtccttatcgatgcttgaatcaacat
cacattgtttgcatttcgtttttgcgtgcaaatatgttatttgcaaagaaggcaaggtaatgtgcttaagagtaaatacaattcgctg
tccatttttgtccaccagtgtgccagaacccgtgcctttagtccttcgaatacatccgaccagtcagcaagcaagtgcatc*ATGG*
TGCTACCGAAGCTGTCCGAACCGTACGCCGTGATGCCGCTTCTACTACGCCTGCAGCG
TTTCGTTGGGCTGTGGGGTGAACGACGCTATCGCTACAAGTTCCGGTTGGCATTTTTA
AGCTTCTGTCTGCTAGTAGTTATTCCGAAGGTTGCCTTCGGCTATCCAGATTTAGAGAC
AATGGTTCGCGGAACAGCTGAGCTGATTTTCGAATGGAACGTACTGTTTGGGATGTTG
CTGTTTTCTCTCAAGCTAGACGACTATGATGATCTGGTGTACCGGTACAAGGACATATC
*AAAGATTG*gtgcgtgataatgattgataaaaggaacctttgagcaactcctatccctttcaag*CTTTCCGTAAGGAC*
GTTCCCTCGCAGATGGGCGACTATCTGGTACGCATCAATCATCGTATCGATCGGTTTTC
CAAGATCTACTGCTGCAGCCATCTGTGTTTGGCCATCTTCTACTGGGTGGCTCCTTCGT
CCAGCACCTACCTAGCGTACCTGGGGGCACGAAACAGATCCGTCCCGGTCGAACATGT
GCTACACCTGGAGGAGGAGCTGTACTGGTTTCACACCCGCGTCTCGCTGGTAGATTAC
TCCATATTCACCGCCATCATGCTGCCTACAATCTTTATGCTAGCGTACTTCGGTGGACT
AAAGCTGCTAACCATCTTCAGCAACGTGAAGTACTGTTCGGCAATGCTCAGGCTTGTG

FIG. 7A

GCGATGAGAATCCAGTTCATGGACCGGCTGGACGAGCGCGAAGCGGAAAAGGAACTGA
*TCGAAATCATCGTCATGCATCAGAAGGCGCTAAA*gtaaggtctgccggtatgttgtggatagaatacattt
ctagctgctttcag*ATGTGTGGAGCTGTTGGAAATCATCTTTCGGTGGGTTTTTCTGGGACAG*
*TTCATACAGTGCGTAATGATCTGGTGCAGCTTGGTTCTGTACGTCGCCGTTACG*gtaacta
aaagcactgtagtgatctgtctgccacaccattcactgctgtgtcttgttttgtcactcttcccag*GGTCTCAGCACAAAAG*
CGGCAAACGTGGGTGTACTGTTTATACTGCTAACAGTGGAAACCTACGGATTCTGCTA
*CTTTGGCAGTGATCTTACCTCGGAGGCAAGTTGTTATTCGCTGA*gtttcagttacttttccgttcccc
tctaaccgtaccacttgtaccatttgtttgagacagagcttgagcgtag*CACGTGCTGCGTACGGTAGCCTCTGG*
TATCGCCGTTCGGTTTCGATTCAACGGAAGCTTCGAATGGTACTGCAGCGTGCCCAGA
AACCGGTCGGCATCTCGGCTGGGAAGTTTTGCTTCGTCGACATTGAGCAGTTTGGCAA
*T*gtatggggagaccttccactgtggcaagaaagattttctttattaatgcatcttttaatttacag*ATGGCAAAAACATCA*
*TACTCGTTCTACATCGTTCTGAAGGATCAATTTTAA*aggggaactcccccacccgaccagacgacggaa
agctaacgatgtgcaattgaatagtcattagtagcgttttgctcgcaaacgaactaacctttgacttttaagttcactacggtgag
gacaaaaatcaataaattaaatcgagaccgttgatgagcaaaagaaaaaaaaatattttactgattttcatttcgttccatcgacta
cataatcataattatatgccacattttattataagttttg

FIG. 7B

Anopheles gambiae odorant receptor 6 partial genomic sequence (SEQ ID NO: 22)

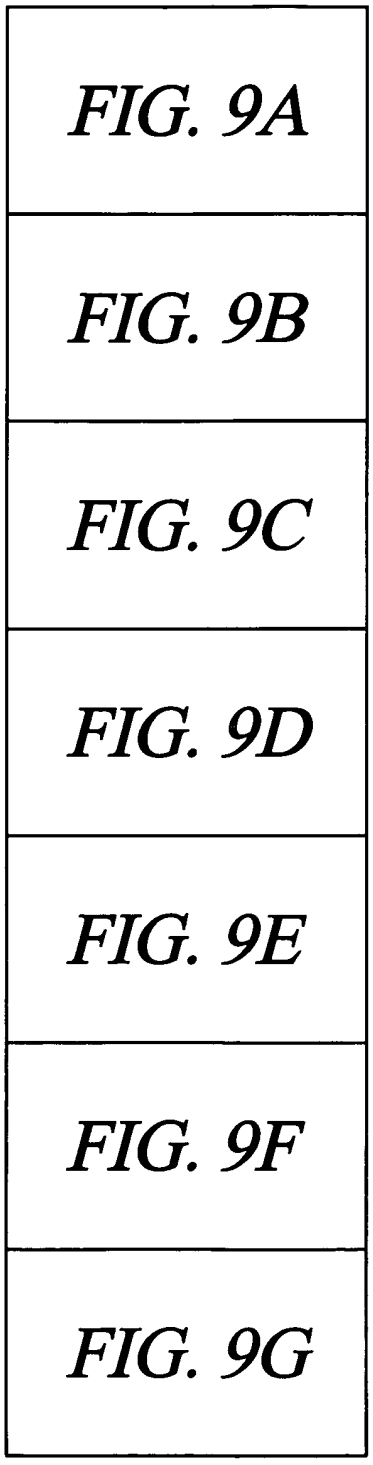
These are the predicted last three exons of another candidate *Anopheles gambiae* odorant receptor.

Predicted Exons: *ITALICIZED*, UNDERLINED AND BOXED was highlighted
Introns: lowercase.

aacacccatcttatcggcaaaattagtatttaccgtttgaaagcggcttcccttcctggctgtttctcactctctctctctgtctctctta
ttgatgccgtatgcgccgcgtgctataggctag*TTATGCTTACCGGATGTTGCGATCGCGCACGTGCTTT*
TCCGCATACGCCAGTGCACACTTGATGGCGGTGGTGATGACGTCTGCTGCGCACCGTT
TTCTGCTCGTGAGTCAGACCTTTTCATTTCCTGCAATATCCTGTTTCTTTCCCGACCCC
ACAGACGGTTAGACGGATATATGCTGGTAAAGTTTGTCCTCTTCATGCTGTGCTTTCTG
*ATCGAGCTGCTGATGCTGTGTGCGTACGGTGAGGATATTGTGGAATCG*gtaaggcaccaggc
ggtgatgagcgagtcgcgagtaattgaagcttttgcttttaaaacacatcagag*CCTTGGGGTGATTGATGCCGCT*
TACGGTTGCGAATGGTACCGGGAAGGGTCGGTGGCGTTCCATCGATCCGTGCTGCAAA
TTATACACCGCAGCCAGCAGTCCGTCATACTGACCGCATGGAAAATTTGGCCCATCCAA
*ATGAGTACTTTCAGTCAG*gtgagttgccaattgattgccgtttgcgttaatatttcagtaagagtgcgctctttcccttag
*ATCCTGCAAGCTTCCTGGTCCTACTTTACCCTCCTGAAGACCGTCTACGGGAATAA*gtaa
gcgcgagagagagagagagagcagtatcgttcacccttggatgaatcaatagatttctaatcatgaaccattgaaaaatgaatca
acattttcgctagttgcacaatattgtaccattctatacagcttcaccacgaccaagcgtttgttgcatcaggaccaaacacgtttcga
caagccgcgtcacctgctggc

enlarged

FIG. 8



Now 9A, 9B, 9C 9D, 9E, 9F, 9G due to enlargement

FIG. 9

Anopheles gambiae odorant receptor 7 genomic sequence (SEQ ID NO: 23)

Features
1. Predicted Exons (7): ALL CAPS, ITALICIZED, UNDERLINED, BOXED
2. Introns (6): lowercase
3. 5' and 3' sequences: lowercase, dotted underlined

was highlighted

ccgcccgggcaggtgacttacgcggtctgacttgctggtgcgctgctttgtacggcaaacggctacacaagcgaatcgaattattttcc
tatcacgctgcgcttaccagcgcctgctggtaggcaaagaatgtgcaaagtttcatttggcttggttcgtctgctttgctgtgaacgtgt
gcacggttgcatcgctaaggtttcggtgtgagccgagaagttgcagatcgaaatctctttgtgtgtgtgtgtgtgtgtgcagtgggaa
gcattgtgtttagtgagaagtgaaaagaaaagtgctgaaaaatgcaagtccagccgaccaagtacgtcggccttcgttgccgacct
gatgccgaacattcgggttgatgcaggccagcggtcaactttctgttccggctacgtcaccggcccgatactgatccgcaaggtgtac
tcctggtggacgctcgccc*ATGGTGCTGATCCAGTTCTTCGCCATCCTCGGCAACCTGGCGACGA*
ACGCGGACGACGTGAACGAGCTGACCGCCAACACGATCACGACCCTGTTCTTCACGCA
CTCGGTCACCAAGTTCATCTACTTTGCGGTCAACTCGGAGAACTTCTACCGGACGCTC
GCCATCTGGAACCAGACCAACACGCACCCGCTGTTTGCCGAATCGGACGCCCGGTACC
ATTCGATTGCGCTCGCCAAGATGCGGAAGCTGCTGGTGCTGGTGATGGCCACCACCGT
*CCTGTCGGTTGTCG*gtatgtgtgtatgtgtgtggccgtttgggaaagtgtctttgcggcagaaccccaatctactgttacgc
ttgactgggtttttgttttttctcggtggagggacgggataaaatatctgaaagaataattgagtcaacccacaggggatgcaag
acatcgcaggcagagagtttgggtttgatttatcaccgcacaccgaatatcttcacggttcataagcttcaccgcggtgaaaaggga
actccccatttccctgttttctttttttttcttcctctcgataaattactcatcgcttttcgttttttttttttttgttgttgcttctttcttctttcatc

FIG. 9A

cctactag*CCTGGGTTACGATAACATTTTTCGGCGAGAGCGTCAAGACTGTGCTCGATAAG*
GCAACCAACGAGACGTACACGGTGGATATACCCCGGCTGCCCATCAAGTCCTGGTATC
CGTGGAATGCAATGAGCGGACCGGCGTACATTTTCTCTTTCATCTACCAGGTACGTTG
*GCGGAAT*gtcctgcgcgtcacagttggcagtcagtgagcggcaacacggcgaaaaaatgggactaaaaccggtcttcacaga
gccaacacattcctacagcaattgcataccttcgggcggtcgggactgggcaatgcagctacaacatcctcgcctaaagttatgcaat
tcgagcgacaaatgttgccgtgttagggctttttgtgataatagtcgtttttttgtcctctcgcttatcaaactctatcaacggaggaaa
tccattttcgctacaatgcctacagctcaagtttcaaggtcaatcgagcgggtggggatcaacttttttattcattttgctaacgcccca
tcaacaaattctatgttctcaatggcaaagattactgcccgcaccaatcgcccaacgaaacggcaaaagaaaagcgacgattatga
agatgtccaaaccattgcccgcccgacgctttatctgatgatttgcgggatggcttttacttgtctgctactttcaggcacaaaaggaa
atgaaaccagcgcaggctcgtttgccggcttgcggaggttcttcaggcactgaggctgagtacttaaatcgaacgattttacgattc
tggatccagttttatgatgtggcctgcattacagtggcaattataccctgatgttcatttcattgcattttgtaagtttgtgctggtaacg
cccgtaacgattaattcttttcaaagagattctttcaaagagattcaaaatgtgtataacaaatgctaacgaatggaccgtacttgg
agggttgcggaaagtaacgttttaaaatattcatcacaatcctctgcaaacttgtgcttaattaattggtgcacaataagtttaaact
gtggcggcagatgtgtcgctgtccgcttccttccttcccagcaagctcgtgcgaaataatttattccatcattttaatacagccgtttgtg
catttaattagcaaagcaatataaaaagcagctaaccatccccattaaaacaaagtgcttccgggcccaattgttatggcggtgga
aagtaatggttttaccagtggaagtgtcctttcccatcgtgggtacttcgcgatattcttgtcttatacaagtgcatacagaaaaaaa
ggacaaatcctccttgctatggtctaaggccagcttcggtaccgcttccgcttcgggatgtcataaagtttgatgggtgtttttaacatt
acttccgctcttaaccacctaatggacttttcatgcttgagctaaagttaaaccagccaccagcggtacgcaccgagccacggttgatt

FIG. 9B

tcggcggcggcctcatccccagttttgcgccaccaatattgccttcattaatctgtaccctcggagcgttagggcccgcggacgagtcct
cgttgtaatgcaccgccatgccacgggacgggataatccgttgggacggcgcgaaagcgactatcgcggacggattggttcgaccg
tgctacaacacattttatgcttcacagatttacttcctgctgttttcgatggtccagagcaacctcgcggatgtcatgttctgctcctggt
tgctgctagcctgcgagcagctgcaacacttgaaggtaggtacggtagcaaacgtggttgtctttacatccgcgtgcagcattatcct
tatcgacgtgtagtgttaacggtaaaagaggaagcgataaaaaagcaacattctctcacaccctcgatctctctttatttttctctctct
ctctctctctctctctctctctctctctctctctctctctctctctctccatctcctcgggcag*GGTATTATGCGATCGTTGATG*
GAGCTTTCGGCCTCGCTGGACACCTACCGGCCCAACTCTTCGCAACTGTTCCGAGCAA
*TTTCAGCCGGTTCCAAATCGGAGCTGATCATCAACGAAG*gtatgtgaaacgtgtgctcgtggcagacg
gactcaaagagagcataacacaatcccctggtagttcatttcaatgaccttaacactcggcaagctaagcgagacagtggggacag
tgagaaagagagaacaagaaaaaaaaccatcatccgtacgacatcatcgctacgtaccggtatttcaggatgaggaaataaaac
gctaggggaatgaaagtgcgacagaatgataaaacaatccccacccaggcccccagcctggacgaacggatgtagtgtgcgaagc
gagcaaaaaaagtcaaataaattgaagtttaaaaatagattttccccgtccatccgtggtggagcgtaaagcccggcggacaactt
cgagcacggcgaccgtgcacagtactgtgccacagttgtagggacggataagctccgttccttttttatccttttttttttggagatttgt
ttgcgttcgcatcgttagacgagcttagtgccgtgttgctctaattgctatttattataaagcgcttccaaatagaagatcggttctctc
catttaatctatcgcgcctgtacgcctgaaactatgcactgtgctgtgaaaccgtcaagctcgagcacgacgaatggcccaccgtacc
acgcccgtggtgcccaaagcgcaacgcgaattgcatgttaacaaaccttgcctaccatccaatccgtgtgaaattgcccgctctcttt
ctctcttttgcgctttcggtgtatcgaacggttttgtccctttttttttacttttgctcttgatctcttgctgtgctcactttcatctcatgttttgc
ctgacggtggtgggttttcgaaaaaagagcgatttcttctgcgtgtgtgtgtggttttttaaataaccgctccaggtcgtgttgaacg
ctgcaggaccgatcggagctagtttattatcagctttagtgtttatcccacccatgccccacatcacgtctgtggagagtgggggaag
cttaagtccaatgtaatttaccgtgtttctgtcgttcgtcaccttcttcgtcgatggagattggtgcggttggcacgataaaagcccact
gcacgttacggaccgagggaaaggtctttttgtaggcctagcaacggtcctcattcaccgcatgggggtgtagctcagatggtagag
cgctcgcttagcatgtgagaggtaccgggatcgatacccggcatctccaacccacacaaaacgtttttttaagaagattttagggaa
gatattaacgcgggtacactgtgctcctctaagttggaagagtagatgagatgatgacaagggagaaggaacatgtgtacgtgttt

FIG. 9C

gatagcaaacacacaaacaacaatatcatctctgataataatctgatgtgtgatgtgtgtattgttgttatgctgcctttgccatct
tgtccctctctctcctgttcaactcctaaaagaattgtttggagtcctctcagttcctcgtaaagatcctttcgagattcttctttccttttt
attatttattccacgagcctctgacataagtagccttccgcttatttccttctccttgcacttgtcagttccgtgtagagcgtcattttgag
gtttacacatttcccaccgacgcctgattgttacattgtcatctacattgctttccgtttaccgttccgccctttttttttaacgctaccaca
g*AAAAGGATCCGGACGTTAAGGACTTTGATCTGAGCGGCATCTACAGCTCGAAGGCGG*
ACTGGGGCGCCCAGTTCCGTGCGCCGTCGACGCTGCAAACGTTCGACGAGAATGGCAG
GAACGGAAATCCGAACGGGCTTACCCGGAAGCAGGAAATGATGGTGCGCAGCGCCATC
*AAGTACTGGGTCGAGCGGCACAAGCACGTTGTACG*gtaggtatggtaatttctaaggtgtggtgtaaag
cctccaggttccatgaaaaagggatactttaccacagtaagagtttgttttgctggacttacattctttggagcattgtttggtgttgtg
ctgaaaccggttgcaatatcgttttgcgaagaaattatgtgtaaagcgtattacaatctcattcctctgttaatctgtaccaattgtgtc
agccccgaccgaaagcaggcctaattcgtaccagaaaaaccacaagctgtttgtaagcatcgatacgcccgaagctttcaatccagc
caaggcgccacctactattgacgtgacttttgcacgttcacactctccctctcccattctttctataaccaatcgtcgctcagccagcat
cgcccggagtgaagttttatttgaacgatatcacccgtatcgatttccactaaacatgcttaaatcgtttcacaaagctcccccaaa
atcccatttcaccaatccaccaatttgaagtccgtcgtcctttgtgtccttgtgtttgtgtgtttgtgtgagctggagacatgggggagt
gagtaaccgaacaacctcttgccgctgcttcacgatatcgaacagcaccaagataagcatccctttttccctagccgatgtctccgata
tctcgattccgcttccagcgaggcaaagaaaaaggcgaactggctgacctcacccggggcgaggaaaaagcgtagggattacgtc
gagcagcacgagttgtgatttcttcttcttctggttccataaatcgctgacggtttccattaccgcctgcggagtgcacacacgtgaag
ggaaagcgaaaacgtttagattccagcagcaacggcagcaccagaagcagcagcagcgcggcaaattgaatcatcctgacgcgat
gagttgtctgggttttcgggtcggtggcttacagcaccacaccatctgctgcagctaatacagctgtaaatttcgttagacatagactt
gattttacaatattacacacacacttacacacacagctatagatttgtcgcttggcgtatggctctgtacggcgtgccgtacatgccgc
gagccgtgttgctgctggttgcgatacggatcacgtccgattcgattcagcctgcgtgtttttggtgaagatccttatcggtgacccact
ttcagtgtgtcgagagcgagggtcactatggcgcctgtcagttggaaagctaggctcgattcaaagggccattgtgccagtgttcttt
ttaagatagcgataagcttttgatcgaaatagtaaatcaaacattgtttctttttttcctattccaaactgttgccaacctcattattacg

FIG. 9D

ttttgcagcgggtgtatagtaaattgcatactttaaggcgtgattttcaaatgtagcgttccgtatgcagaaacgccatggattatgc
aatttaaacaatgctgcttccttaacattcaaataacggcttattaaggaacttttgtgcaatttgtttttaacagcaaatagttagc
tcagaacgatcacatttagtatcgcttcaacaaagaactcttttaaacacacaatttgtaatgccattccctcgagaaagtttcttgtc
agtcctcctctgcatcacagcaacaaccaaacctgctcatgtttcctgctcgtttcctagctgttttgaacgttatttccgattcctgtgct
tgcccgcttttcttacaatcaaccacaatggttcagatttcgctcttattttattgacccactgctttcgtgctgaagcccgtggaaacaa
tgcgccaagctcagcatccagccatgcatgtaaaatgagccacgcgacagattttagacatcgctttcgctctgcaccggaggtggtt
ttattcttgtttccgattcccacgtccattcgtcctgggtccgtccgccgggcccgaaaccgtaagccgtgcggggaattacgcaatcga
aacgagccagaaaatgagcacgccaaatgcaaagaaaatccccttttgagtggtgctcctgccaccactcatctccccaactggtgg
gtgaaaaaccttgtgcgccccttctctttccagaaaaaaaacgcctcgctcgcacaaaaacatgctcgcccggtgaagctgcgtatgt
cgcagaagctcaaaccaacgccgccagcaagcatcaacaatttctattcaaacacccaacgcagcgcccaaaccgggtgcactgta
ctcagtagcgaagatgctcagattgtcccgtgcgctgctttcgatgcccgtttcggagcgggaagccatcgcttgccaacgttggcgat
gtcttttagccgtggatttgaattttctgaatatcacaggcgggcgcggtttgcctgcaaggttgttgcttcccacacgagcattgcttt
ccgtaccgcggtggggcgagttttcaacgcaaccttctacaagcaacgccacaacgcctgggagcgatatttaacagaaacaagaa
catcccgaacttcagcacatgccgtgatttgcctgttggaaaagcttttgtgagcgtgtgagttgaacgagctctattttcccagcgat
gggtggcatttgtgtggcatgctatcgtcagcttttcttgaatctttacctctccattcgcctccattagtacacgcgtatggaaaatgg
gtgcaacggatcagaacggattttccgcgacagacttaataaagggaaagcaacgcgtttttttgcatgtgtagtgtttatgagcttt
atgccgttactttgcaattaaaaatagcaaaaaataacagtttttttttgtaagcggattacaaagaatgtatcagaatattacgtg
aaacattcatttcatgctgttaacgctcaaatagaatagttttgtaacacggattgcataccttgccggtatcggttacattttcgccta
acagtatgcaatctgtttagctttgttgtttaatgactgcgttggtagtacaatatttatttacaccgcgtaatttatctcacaaattgc
aaaaaaatgtcaatctgtatcgattattcacacaaatcagatcccggaaccagtgtagcccaatgtgctcttattgaattaccacga
acaaatcaacctgatgcccgggtccgttggcaaacagcttgcgccgaagccgctcagtgtttcgtgcactaccgtgctgccattttgct
gccctcatcgaacagataaacagaagggcaactcttgtgagcatcgcaatgcccgtctgaagttccgtcgaaaatgggcctaaattc
aatttgacgcatttacccgcgaacaattgcgcgaaggctgtcaagtgtgttccacgaactgcgacaacaagcacacacacaaacac

FIG. 9E

aaatgttatcgtttcggcatgtttctcggtacaaagcgtgtggcgctatgtggcatgccgattcccagacagagtgatcgatagtaaa
tgtagcctatccggtagcattcaatttccttttctatcctcgcaaacaaagcccattctggggaggcgtggtgaagctttcaaaggcat
tgtgaaacaaatgtcctggttcggagggatgctggggaaagcaaacacggtgccgccatcgctgctaccgtcaatcgatcatgcatg
atgtgattaatatttgtgttattcacctgcgtatctatgcgtccgtcgtgtcgttcggatttccggaagtcaaggaaaaagcgactcca
tttgggattggtttttgcagcgaaaaatcaaaacattcgcacaaaaccgtcctccatttcaaatgcctacacttgtcactgtatatctct
ctttctctcgttttgccacgttgcag*TCTCGTTTCAGCAATCGGAGATACGTACGGTCCTGCCCTGCT*
GCTACACATGCTGACCTCCACCATCAAGCTGACGCTGCTCGCCTACCAGGCAACGAAA
ATCGACGGTGTCAACGTGTACGGATTGACCGTAATCGGATATTTGTGCTACGCGTTGG
*CTCAGGTTTTCCTGTTTTGCATCTTTGGCAATCGGCTCATCGAGGAG*gtacgtgcgctcggcgtg
ttgccgtgggaaagcattctccctgccccatatcgcttcattctcccagatcacacatttgcatcacaaagccagcacacttttgcttcg
ccgctgccatctcggcttctgaatgttttcacttctcccatacttctcccgtgcag*AGCTCATCCGTGATGAAGGCGGC*
CTATTCCTGCCACTGGTACGACGGGTCCGAGGAGGCAAAAACCTTCGTCCAGATCGTT
TGTCAGCAGTGCCAGAAGGCGATGACTATTTCCGGAGCCAAGTTTTTCACCGTTTCGC
*TCGATCTGTTTGCTTCG*gtaagtgtagcctggtggctggcacagaacaggctggcaaaacagggactttggctctagc
ctgatgggtggtatatgtgtgtctatttttgctaccattctcgcatccttcctttccag*GTTCTTGGAGCCGTTGTCAC*
*CTACTTCATGGTGCTGGTGCAGCTGAAGTAA*acagccgtggcccggaaggatgtgtttttttcgctcgttcg....
gttgtttgtttgtgcacactttctcttggacattttctctactgcaaaggtttaacaaacagcaacaacaaataatcccaagttttctttt...
acagatctttgcaaaatgattagattttaatagattaacagtgcttgattatctgtcctgtagcaaccggggctgaagaacgttgatt...
tggtaaaagtacaaaagggacgttggaaattgaaccaccagaagagtgatatttatgcaaagctcaccaagggaaatctatgtat...
gtgtgatttgcgctcatcaagcactgtatgtgcctttcaactagtgcagcaataaagagtacaaatgtttcttagcgcaccgtacattg...
tcgtttcggcgttttaaccgttgttgataatacacaaaagatgataaaaataaataataacaaaatgttaatatgagtaagtacta...
aatagagaaatcgttttagtatgatcatacctccaatcatttgtttgaaattaactttaattttaactcaaattaaaccgatgttttact...
ttctgtgagaattattgtggaagaacttaatggaagtataattaaattgattgctaactttatgcgtttttcaatttacgaacgctagt...

FIG. 9F

cttcaaacatcgcttcaaaagtattactaccacattattcatttacttatagttatatttattgcctcttcatctttccatggccagaact
actgcagaaaagcttctttttgctcgctttccgatggttggttggacgaagttggtaacaaacggcaagcaattagcataaactatt
ttcgcatcgagatggaaatgaatgtaccactagaaccgagtgaaatgaattacttttcaacttgcacgccaaaaccattatctaaag
tacgcacaacttaaaaacaaaaccccaaattgtcgtccacccttcattccactttcttgctacactttccgaccgagttctgtagcgccag
cagcaaaaaaatacatataaaaccttcatcactcaagctgtatcgagccagcgtgggttgtgtttgactgtgctgtgaaagaaaga
agaaaaaaaaaaacacttccacgggaagctagcaattggaaatgcataaattaaccggaagaaattcgcaaaaccccgcaccgac
gtaccgcaccgcatccgtaccgataccggaacaaacggtgtgcgcgaaagaatccgctagcagccccactggcacgggtatttgctt
ttggttctgtgtttttcttccactggtttgggtgcctgggcgaaggctagctcggctactttcccggggccgcaattttctgcagcccaag
gcggcgtgctcgtggggccaaaagaat

FIG. 9G